Rec'd PCT/PTO 11 JAN 2005
10/520882

Transgenic plants used as a bioreactor system

BACKGROUND OF THE INVENTION

FIELD OF THE INVENTION

The present invention relates generally to the use of plants as bioreactors for the production of molecules having useful properties such as *inter alia* polymers, metabolites, proteins, pharmaceuticals and nutraceuticals. More particularly, the present invention contemplates the use of grasses, and even more particularly C4 grasses, such as sugar-cane, for the production of a range of compounds such as, for example, polyhydroxyalkanoates, pHBA, vanillin, indigo, adipic acid, 2-phenylethanol, 1,3-propanediol, sorbitol, fructan polymers and lactic acid as well as other products including, *inter alia*, other plastics, silks, carbohydrates, therapeutic and nutraceutic proteins and antibodies. The present invention further extends to transgenic plants and, in particular, transgenic C4 grass plants, capable of producing the compounds noted above and other products, and to methods for generating such plants. The ability to utilize the high growth rate and efficient carbon fixation of C4 grasses is advantageous, in that it obviates the significant growth penalties observed in other plants, and results in high yields of desired product without necessarily causing concomitant deleterious effects on individual plants. In addition, the C4 grass, sugarcane, is particularly advantageous, as in addition to the features common to all C4 grasses, this plant accumulates sucrose. This sucrose store provides a ready supply of carbon based compounds and energy which may further obviate any deleterious effects on the growth of the plant associated with the production of the product. The present invention provides, therefore, a bioreactor system comprising a genetically modified plant designed to produce particular metabolic or biosynthetic products of interest.

DESCRIPTION OF THE PRIOR ART

Reference to any prior art in this specification is not, and should not be taken as, an acknowledgment or any form of suggestion that this prior art forms part of the common general knowledge in any country.

Modern techniques of biotechnology are driving a new revolution that promises both scientific and financial gains for a range of industries. One difficulty, however, is the large financial cost of establishing sufficient infrastructure to generate recombinant products or to generate the products resulting from recombinant processes. Alternative, more cost effective systems are required to assist the generation of large amounts of product resulting from recombinant processes.

For agricultural industries, the generation of genetically engineered plants enables plants to be quickly developed with desired traits such as resistance to pathogen infestation. However, plants can also be used to produce a wide range of compounds not normally produced within the plant, thereby providing a source of renewable raw materials for the manufacturing, energy and pharmaceutical industries.

This endeavour is aided by the fact that plants, animals, insects, bacteria, fungi and even viruses have evolved in a wide range of different habitats and, hence, produce a remarkable array of compounds which allow them to survive and thrive under very varied environmental conditions. It is estimated that up to 100,000 unique compounds exist in the plant kingdom alone. In the future, genes and even entire genetic pathways may become available from different sources to assist in the manufacture of a wide range of commercial products.

Traditional chemical industries are increasingly looking towards biological systems for the production of bulk and fine chemicals. Biological processes offer numerous advantages over chemical processes, including the elimination of complicated and difficult high pressure and high temperature reactions, the use of aqueous systems rather than organic

solvents, high degrees of product stereo-specificity, a capacity for highly complex synthesis and comparatively simple scale-up. The use of biological processes is not a new phenomenon, as many fine chemicals (e.g. enzymes, antibiotics) and bulk chemicals (e.g. ethanol, amino acids, citric acid, lactic acid) are produced effectively in microbial systems. Advances in molecular biology and genomics have enabled an expansion of the available product range, the transfer of production systems to microbes with desirable production traits, and significantly increased yields. Nevertheless, inherent limitations remain, in that the raw materials (e.g. molasses, sucrose, or high fructose corn syrup) and scaled up fermentation processes are relatively expensive.

By contrast, genetically modified plants should not require any more raw materials than are already required by their non-transformed counterparts and have the potential to provide a low cost of production per tonne of biomass, when compared with fermentation methods. Thus, if reasonable product yields could be achieved in plants, and if these products could be extracted at reasonable cost, the potential for chemical production in plants would be extremely high.

The polyhydroxyalkanoate (PHA), poly-(D-3-hydroxybutyrate) (PHB), is a thermoplastic with physical properties akin to polypropylene. Both PHB and polypropylene are water insoluble, exhibit good gas-barrier properties and possess similar melting points, degrees of crystallinity and glass-rubber transition temperatures (De Koning, *Can. J. Micro. 41(1):* 303-309, 1995), although PHB is more resistant to UV radiation. Moreover, unlike polypropylene, PHB is rapidly degraded by numerous bacteria and fungi under composting conditions (Jendrossek *et al.*, *App. Micro. Biotech. 46:* 451-463, 1996; Mergaert and Swings, *Indust. Micro. Biotech. 17:* 463-469, 1996).

Cost is the major reason why PHA produced by bacterial fermentation cannot compete with conventional plastics production methodologies. Major contributors to the cost of PHA are substrate cost, energy consumption during fermentation, disposal of waste product, and the cost of constructing and maintaining plant and machinery. The use of transgenic plants for PHA production, however, has the potential to either eliminate or

drastically reduce these costs since atmospheric carbon dioxide would be the substrate and energy would be derived from sunlight. Operating costs would be no more than what is incurred in ordinary agricultural practices. Waste products are the same as for a non-transgenic crop. This makes plants an attractive potential alternative to bacterial fermentation.

The production of PHB has been most closely studied in the bacterium *Ralstonia eutropha* (formerly *Alcaligenes eutrophus*), which accumulates PHB at up to 80% of its cell dry-weight (Steinbüchel and Schlegel, *Mol. Micro. 5:* 535-542, 1991). The PHB biosynthetic pathway within *R. eutropha* is well known and consists of three steps catalyzed by the three enzymes 3-ketothiolase, acetoacetyl-CoA reductase and PHB synthase, respectively. For large-scale industrial production of PHB, bacterial fermentation is economically and environmentally less favourable than the corresponding production of petrochemically derived plastics like polypropylene (Lee, *Trends Biotech. 14:* 431-438, 1996; Gerngross, *Nature Biotech. 17:* 541-544, 1999). Hence, since the glucose supplied to PHB-producing bacteria is derived from plants, it would be advantageous to be able to produce PHB in plants directly.

Attempts to achieve this goal have been unsuccessful, due largely to the significant added burden placed upon individual plants by requiring them to produce this macromolecule, resulting in a severe reduction in plant growth and infertility (Poirier *et al.*, *Science 256:* 520-523, 1992; Bohmert *et al.*, *Planta 211:* 841-845, 2000).

In accordance with the present invention, an efficient bioreactor system is developed in plants and in particular *Saccharum sp.* such as sugarcane.

SUMMARY OF THE INVENTION

Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element or integer or group of elements or integers but not the exclusion of any other element or integer or group of elements or integers.

Nucleotide and amino acid sequences are referred to by a sequence identifier number (SEQ ID NO:). The SEQ ID NOs: correspond numerically to the sequence identifiers <400>1 (SEQ ID NO:1), <400>2 (SEQ ID NO:2), etc. A summary of the sequence identifiers is provided in Table 1. A sequence listing is provided after the claims.

The present invention provides a plant-derived bioreactor system. Although previous attempts to effect the manufacture of useful industrial and other products in plants have not been overly successful, it has been determined in accordance with the present invention that this was due to the extra load placed on individual plants which resulted in deleterious growth effects. C4 grasses, which have particularly efficient mechanisms for the assimilation of carbon, are identified as having a high growth rate and high accumulation of biomass making them useful as bioreactors for the production of a wide range of products. Furthermore, the C4 grass, sugarcane, is particularly useful as this plant stores sugars in dimeric and/or polymeric forms. These stores may be utilized when needed such as, for example, for rapid vegetative growth, orfor energy during times of significant environmental stress. The store of carbohydrate is identified in accordance with the present invention as providing a ready supply of precursor for many metabolic pathways, and utilisation of this store does not stress the producing plant. Therefore, the present invention is predicated, in part, on the identification of a subset of plants, namely the C4 grasses (particularly sugarcane), as useful bioreactors on the basis of their high carbon assimilation rate, rapid growth, high biomass production and large carbohydrate store, such as is found in the stem.

The present inventors have capitalized on the potential of a crop that has a highly efficient

C4 carbon assimilation mechanism, a rapid growth rate and naturally harbours large quantities of sucrose in its stems, thereby having the ideal properties of a bioreactor. The instant inventors have developed a means to engineer this crop so as to effectively accumulate significant quantities of a product without significant decreases in biomass or growth rate. In so doing, the plants contemplated herein permit the manufacture of products such as biodegradable plastics, vanillin, indigo, adipic acid, 2-phenylethanol, 1,3-propanediol, sorbitol, fructan polymers and lactic acid as well as therapeutic, nutraceutic and diagnostic agents without incurring the previously observed deleterious effects on growth and viability.

Accordingly, one aspect of the present invention provides a method for generating a plant-based bioreactor system, said method comprising selecting a plant having a high efficiency carbon assimilation mechanism, rapid growth rate and/or high biomass production and/or reserves of metabolites or having a capacity to generate such reserves and/or which possess metabolic and/or biosynthetic pathways useful in the manufacture of a product of interest or a precursor form thereof; genetically modifying cells of the plant to enable access to said metabolites and/or metabolic or biosynthetic pathways; and then regenerating a genetically modified plant from said cells.

The present invention is particularly directed to C4 grasses, however, other non-grass C4 plants such as woody or herbaceous plants which utilise the C4 pathway are also contemplated by, and are within the scope of, the present invention.

In a preferred embodiment of the present invention, the subject plant is a C4 grass. In an even more preferred embodiment the plant is sugarcane. As used herein, the term sugarcane is to be understood to include, *inter alia,* plants of the *Saccharum* genus, incl. *S. robustum*, *S. offinarum*, and *S. spontaneum* and hybrid *Saccharum* sp., incl. modern sugarcane cultivars.

The preferred compounds to be produced by the plant bioreactor include: vanillin, sorbitol, polyhydroxyalkanoates (PHA) such as poly-(D-3-hydroxybutyrate) (PHB), indigo, fructan,

lactic acid, adipic acid, 1,3-propanediol, 2-phenylethanol and pHBA. However, the present invention also extends to the use of C4 grasses as bioreactors to generate a compounds such as therapeutics, nutrapharmaceuticals, diagnostic agents including, for example, single chain antibodies, industrial enzymes and the like.

The present invention contemplates, therefore, a method for producing a product of interest including a product or intermediate of a biosynthetic or metabolic pathway in a C4 grass, said method comprising expressing one or more genetic sequences which encode one or more enzymes or proteins required for the production of the product or intermediate or a homolog or precursor thereof or which induces gene silencing of genetic material which encodes an enzyme or protein in a biosynthetic or metabolic pathway in cell of a C4 grass plant such that the product or intermediate accumulates in the cytosol, storage vacuole, plastid or non-plastid organelles of the cell, or accumulates in the juice or vascular fluid of the plant.

The present invention therefore provides for the production of a product in a C4 grass wherein product accumulation is at least in part predicated on the direct activation or inhibition (including down-regulation) of an enzyme in a biosynthetic or metabolic pathway by the administration to the plant of an enzyme inhibitor or activator. Reference to an "enzyme inhibitor or activator" includes genetic materials which, for example, induce (post-transcriptional or transcriptional gene silencing of a structural gene or positive or negative regulator gene.

In one preferred embodiment, increased accumulation of a product or intermediate from a biosynthetic or metabolic pathway is a result of inhibition of one or more biosynthetic or metabolic enzymes and optionally re-directing metabolites down another biosynthetic or metabolic pathway.

In another preferred embodiment, the present invention contemplates the use of sugarcane as a bioreactor. Therefore, alteration to the gene expression profile of sugarcane to effect the production of an endogenous metabolite at an increased level, or to produce any

heterologous metabolite is within the scope of the present invention. Accordingly, induction or supression of any biosythetic genes in sugarcane, such as described herein, is to be considered within the scope of the present invention.

Typically, the production of one or more metabolites or heterologous proteins, polypeptides or peptides in a plant is achieved by expression of a nucleic acid molecule encoding the metabolite or protein, polypeptide or peptide of interest. Any nucleic acid which encodes a protein, polypeptide or peptide of interest is contemplated by the present invention. However, preferred nucleic acids include those encoding:

(i) vanillin biosynthetic enzymes, including 3-dehydroshikimate dehydratase, catechol-o-methyltransferase, aryl aldehyde dehdrogenase, feruloyl-CoA synthetase, enoyl-CoA hydratase/aldolase;

(ii) sorbitol biosynthetic enzymes, including glucose/fructose oxidoreductase;

(iii) PHA biosynthetic enzymes, including 3-ketothiolase, acetoacetyl-CoA reductase, PHA synthase, enoyl hydratase, 3-hydroxyacyl-acyl carrier protein:CoA tranferase;

(iv) indigo biosynthetic enzymes, including tryptophanase, L-tryptophan indole lyase, napthalene dioxygenase, *R. eutrophica bec* gene product;

(v) fructan biosynthetic enzymes, including fructosyltransferases and levansucrases;

(vi) lactic acid biosynthetic enzymes, including lactate dehydrogenase;

(vii) adipic acid biosynthetic enzymes, including 3-dehydroshikimate dehydratase, protocatechuate decarboxylase and catechol 1,2-dioxygenase;

(viii) petroselinic acid biosynthetic enzymes, including 3-ketoacyl-ACP synthase;

(ix) 1,3-propanediol biosynthetic enzymes including glycerol dehydratase, 1,3-propanediol oxidoreductase, glycerol-3-phosphate dehydrogenase, and glycerol-3-phosphatase; and/or

(x) 2-phenylethanol biosynthetic enzymes including aromatic-L-amino acid decarboxylase, 2-phenylethylamine oxidase and aryl alcohol dehydrogenase.

(xi) pHBA biosynthetic enzymes including 4-hydroxycinnamoyl-CoA hydratas/lyase (HCHL) and chorismate pyruvate lyase (CPL).

Any of a number of products may be produced according to the present invention. Examples of compounds that may be produced *via* metabolic engineering of a subject plant include: vanillin (4-hydroxy-3-methoxybenzaldehyde); sorbitol; PHAs; indigo; fructan; lactic acid (2-hydroxypropanoic Acid); adipic acid; 1,3 propanediol, 2-phenylethanol and pHBA. These compounds, however, are only examplary, and the present invention is predicated on the use of C4 plants as bioreactors for any compound that can be synthesised in the plant. Accordingly, the present invention is not limited to any one product or method for producing the product.

In a particularly preferred embodiment, the present invention contemplates a method for accumulating polymers comprising one or more species of hydroxyalkanoic acid monomer in a C4 grass, said method comprising expressing one or more genetic sequences which encode enzymes required for the production of the polymers or a homolog or precursor thereof in a cell of a C4 grass such that PHA polymers accumulate in the cytosol, storage vacuole or plastid or non-plastid organelle of said cell.

The present invention further contemplates a method for generating a plant which produces PHAs, said method comprising introducing into cells of said plant a genetic sequence

comprising:-

(i) a nucleotide sequence encoding a phaA or homolog thereof;

(ii) a nucleotide sequence encoding phaB or homolog thereof;

(iii) a nucleotide sequence encoding phaC or homolog thereof;

(iv) a nucleotide sequence encoding phaC1 or homolog thereof;

(v) a nucleotide sequence encoding phaG or homolog thereof;

(vi) a nucleotide sequence encoding phaJ or homolog thereof

(vii) SEQ ID NO:1 or SEQ ID NO:3 or SEQ ID NO:10 or SEQ ID NO:12 or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to SEQ ID NO:1 or SEQ ID NO:3 or SEQ ID NO:10 or SEQ ID NO:12 or a complementary form thereof under low stringency conditions;

(viii) SEQ ID NO:4 or SEQ ID NO:6 or SEQ ID NO:13 or SEQ ID NO:15 or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to SEQ ID NO:4 or SEQ ID NO:6 or SEQ ID NO:13 or SEQ ID NO:15 or a complementary form thereof under low stringency conditions;

(ix) SEQ ID NO:7 or SEQ ID NO:9 or SEQ ID NO:16 or SEQ ID NO:18 or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to SEQ ID NO:7 or SEQ ID NO:9 or SEQ ID NO:16 or SEQ ID NO:18 or a complementary form thereof under low stringency conditions;

(x) SEQ ID NO:19 or SEQ ID NO:21 or SEQ ID NO:22 or SEQ ID NO:24 or SEQ ID NO:25 or SEQ ID NO:27 or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to SEQ ID NO:19 or SEQ ID NO:21 or SEQ ID NO:22 or SEQ ID NO:24 or SEQ ID NO:25 or SEQ ID NO:27 or a complementary form thereof under low stringency conditions;

(xi) SEQ ID NO:28 or SEQ ID NO:30 or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to SEQ ID NO:28 or SEQ ID NO:30 or a complementary form thereof under low stringency conditions;

(xii) SEQ ID NO:31 or SEQ ID NO:33 or a nucleotide sequence having at least 60% similarity thereto or capable of hybridizing to SEQ ID NO:31 or SEQ ID NO:33 or a complementary form thereof under low stringency conditions;

and then regenerating a plant from said cells.

A convenient C4 grass for use in the present invention is sugarcane. Sugarcane has certain advantages, which make it a useful crop for use as a bioreactor including, *inter alia*, its efficient carbon fixation, high biomass accumulation, rapid growth and natural ability to accumulate large quantities of sucrose. Moreover, it achieves this very efficiently by collecting solar radiation and converting it into a carbon sink (i.e. sucrose). Sugarcane is also a hardy crop, is relatively easy to grow and provides a large biomass capability. A micropropagation system is already available and an industry infrastructure in existence.

In another embodiment, the present invention provides a method for generating a plant which produces vanillin or a precursor thereof, said method comprising introducing into cells of said plant a genetic sequence comprising at least one of the following:-

(i) a nucleotide sequence encoding a 3-dehydroshikimate dehydratase;

(ii) a nucleotide sequence encoding catechol-o-methyltransferase;

(iii) a nucleotide sequence encoding aryl aldehyde dehydrogenase;

(iv) a nucleotide sequence encoding feruloyl-CoA synthetase;

(v) a nucleotide sequence encoding enoyl-CoA hydratase ;

(vi) a nucleotide sequence encoding enoyl-CoA aldolase; and/or

(vii) a nucleotide sequence encoding a homolog of any one of (i) through (vi)

and then regenerating a plant from said cells.

Another aspect of the present invention contemplates a method for producing sorbitol in a C4 grass, said method comprising expressing one or more genetic sequences encoding a glucose-fructose oxidoreductase, in cells of a C4 grass such that sorbitol accumulates anywhere in the cell or extracellular matrix of the plant.

In another embodiment, the present invention is directed to a method for generating a plant which produces indigo or a precursor thereof, said method comprising introducing into cells of said plant a genetic sequence comprising at least one of the following:-

(i) a nucleotide sequence encoding genetic sequences encoding tryptophanase;

(ii) a nucleotide sequence encoding L-tryptophan indole lyase;

(iii) a nucleotide sequence encoding napthalene dioxygenase;

(iv) a nucleotide sequence comprising the *R. eutropha bec* gene;

(v) the nucleotide sequence set forth in Genbank accession number D14279, or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to Genbank D14279 under low stringency conditions.

(vi) the nucleotide sequence set forth in Genbank accession number M83949, or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to Genbank M83949 under low stringency conditions.

(vii) the nucleotide sequence set forth in Genbank accession number AF306552, or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to Genbank AF306552 under low stringency conditions.

and then regenerating a plant from said cells.

In another embodiment, the present invention relates to a method for generating a C4 grass plant which produces a fructan or a precursor thereof, said method comprising introducing into cells of said plant a genetic sequence comprising at least one of the following:-

(i) a nucleotide sequence encoding a fructosyltransferase

(ii) a nucleotide sequence encoding a levansucrase;

(iii) the nucleotide sequence set forth in Genbank accession number AY150365, or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to Genbank AY150365 under low

stringency conditions.

and then regenerating a plant from said cells.

In another embodiment, the present invention contemplates a method for generating a plant which produces lactic acid or a precursor thereof, said method comprising introducing into cells of said plant a genetic sequence encoding lactate dehydrogenase and then regenerating a plant from said cells.

In another embodiment, the present invention contemplates a method for generating a plant which produces adipic acid or a precursor thereof, said method comprising introducing into cells of said plant a genetic sequence comprising at least one of the following:-

(i) a nucleotide sequence encoding a 3-dehydroshikimate dehydratase and/or;

(ii) a nucleotide sequence encoding protochatechuate decarboxylase;

(iii) a nucleotide sequence encoding catechol 1,2-dioxygenase;

(iv) a nucleotide sequence encoding 3-ketoacyl-ACP synthase; and/or

(v) a nucleotide sequence encoding a homolog of any one of (i) though (iv)

and then regenerating a plant from said cells.

In another embodiment, the present invention contemplates a method for generating a plant which produces 1,3-propanediol or a precursor thereof, said method comprising introducing into cells of said plant a genetic sequence comprising at least one of the following:-

(i) a nucleotide sequence encoding a glycerol dehydratase and/or;

(ii) a nucleotide sequence comprising the *dhaB* gene from *Klebsiella pneumoniae*, or a homolg thereof;

(iii) a nucleotide sequence encoding 1,3-propanediol oxidoreductase;

(iv) a nucleotide sequence comprising the *dhaT* gene from *Klebsiella pneumoniae* or homolg thereof;

(v) a nucleotide sequence encoding glycerol-3-phosphate dehydrogenase;

(vi) a nucleotide sequence encoding glycerol-3-phosphatase; and/or

(vi) a nucleotide sequence encoding a homolog of any one of (i) though (vi)

and then regenerating a plant from said cells.

In another embodiment, the present invention contemplates a method for generating a plant which produces 2-phenylethanol or a precursor thereof, said method comprising introducing into cells of said plant a genetic sequence comprising at least one of the following:-

(i) a nucleotide sequence encoding a aromatic-L-amino acid decarboxylase;

(ii) a nucleotide sequence encoding 2-phenylethylamine oxidase;

(iii) a nucleotide sequence encoding aryl-alcohol dehydrogenase; and/or

(iv) a nucleotide sequence encoding a homolog of any one of (i) though (iii)

and then regenerating a plant from said cells.

In another aspect, the present invention contemplates a method for generating a plant which produces pHBA or a precursor thereof, said method comprising introducing into cells of said plant a genetic sequence comprising at least one of the following:-

(i) a nucleotide sequence encoding hydroxycinnamoyl-CoA hydratase/lyase;

(ii) a nucleotide sequence encoding chorismate pyruvate lyase ;

(iii) a nucleotide sequence comprising the *ubiC* gene from *E. coli*, or a homolg thereof; and/or

(iv) a nucleotide sequence comprising the *HCHL* gene from *Pseudomonas fluorescens* or homolg thereof;

and then regenerating a plant from said cells.

Preferably, the plant is a C4 grass and, in a particularly preferred embodiment, sugarcane.

A further aspect of the present invention provides a transfected or transformed cell, tissue, or organ from a C4 grass, which comprises a nucleotide sequence encoding one or more enzymes required for the production of a useful product as well as severed or cut parts of a genetically modified plant including stem, flower, seed or other reproductive parts.

Accordingly, another aspect of the present invention provides a genetically modified C4 grass having cells carrying one or more genetic sequences such that one of the following products in the cytosol, storage vacuole, non-plastid organelle or extracellular matrix of said cells:

(i) polyhydroxy-alkanoate polymers

(ii) vanillin

(iii) sorbitol

(iv) indigo

(v) fructans

(vi) lactic acid

(vii) adipic acid

(viii) 1,3-propanediol

(ix) 2-phenylethanol

(x) pHBA

The present invention extends to parts of plants tissue including leaves, stems, vascular bundles, bark, reproductive material, roots and any extracted liquid ("juice") from said plant.

In order to direct product accumulation to a desired sub-cellular location, particular specific "target sequences" may be incorporated into the genetic constructs described above.

A target sequence includes a signal sequence such as a signal sequence to direct the protein to a plastid, vacuole, mitochondrion, peroxisome or ontologically related organelles, or other appropriate organ or tissue.

The plants of the present invention may also be further "tagged" with a reporter that identifies the plant as a plant bioreactor. Any number of physiological or genetic tags

would be suitable, and readily identified by one of skill in the art.

Accordingly, the present invention contemplates a plant suitable for use as a bioreactor that has been tagged with a genetic sequence which encodes or comprises a genotypic or phenotypic feature that allows differentiation of the plant bioreactor from a wild-type plant, or which identifies the plant as a propietary plant.

The plant-based bioreactor system of the present invention is useful in enabling the production of molecules such as PHAs, pHBA, vanillin, sorbitol, indigo, fructans, lactic acid, adipic acid, 1,3-propanediol, 2-phenylethanol, *inter alia*, by a number of different parties such as different commercial entities. The present invention extends, therefore, to a data processing system to monitor the use of the plants and/or the production of target molecules.

Accordingly, another aspect of the present invention contemplates a method for generating a target molecule in a sucrose-accumulating plant, said method comprising:-

(i) providing a plant or cells of a plant to a party; and

(ii) permitting the party to generate and harvest molecules from said plant or cells of said plant receiving and processing data from said party.

The data received from the party includes, for example, numbers of plants grown and/or harvested, the types of genetic constructs introduced into the cells and/or income received from sale of the products.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is a schematic representation showing the four different biosynthetic strategies employed by bacteria that use PHAs as an energy source.

Figure 2 is a diagrammatic representation of the biosynthetic pathway of PHB in the bacterium *R. eutropha*. This pathway comprises three steps catalyzed, respectively, by three enzymes: first, two molecules of acetyl-CoA are condensed to acetoacetyl-CoA by 3-ketothiolase (encoded by *pha*A); secondly, acetoacetyl-CoA is reduced to D-3-hydroxybutyryl-CoA by acetoacetyl-CoA reductase (*pha*B); and thirdly, D-3-hydroxy-butyryl-CoA is polymerized to PHB by PHB synthase (*pha*C).

Figure 3 is a diagrammatic representation of the flow of carbon in sub-cellular compartments within a sugarcane cell.

Figure 4 shows a graphical representation of the detection of PHB in chloroplasts of transgenic sugarcane. A-C: Detection of PHB by HPLC. A: WT sugarcane (s/c) -ve control; B: sample in A spiked with PHB; C: plastid-targeted, PHB +ve s/c line. Arrows: elution point of crotonic acid (PHB breakdown product). Insert in C shows that the peak at 30 min in C has the same spectrum as crotonic acid. D-F: Detection of PHB granules by transmission electron microscopy. D: Chloroplast (c/p) of mesophyll cell of PHB +ve Arabidopsis control; E: c/p of mesophyll cell of PHB +ve s/c line; F: c/p of bundle-sheath cell of same line in E. Scale bars = 200 nm.

Figure 5 illustrates the agronomic performance of PHB-producing sugarcane lines. Four transgenic sugarcane lines expressing the PHB biosynthesis genes of *R. eutropha* (filled bars) were grown for 3 months in a randomised glasshouse plot, together with GFP expressing (open bars) and tissue-culture-regenerated WT (hatched bars) plants as controls. PHB content in lamina from the tips of mature leaves was quantified by HPLC analysis. Data are the mean ± SE ($n = 3$). DW = dry-weight.

Figure 6 shows the affect of PHB production on sugarcane sugar accumulation. Mature- (A-D) and intermediate-aged (E-H) stem internodes from the PHB producing (solid bars), GFP expressing (open bars) and WT (hatched bars) sugarcane plants in Figure 5 were assayed for sucrose, glucose and fructose concentrations. Data are the mean ± SE ($n = 3$). DW = dry-weight.

Figure 7 is a graphical representation showing the distribution of PHB throughout a PHB-producing sugarcane line. The distribution of PHB throughout transgenic sugarcane line PHB3 in Figure 5 was determined by HPLC analysis. Samples were taken from: lamina of the tip, midpoint and base of young, intermediate and mature leaves; rind + pith of young, intermediate and mature stem internodes; and roots. Data are the mean % of leaf DW as PHB ± SE ($n = 3$). ND = not detected.

Figure 8 is a graphical representation showing the indigo biosynthetic pathway.

Figure 9 is a graphical representation showing adipic acid biosynthesis from glucose *via* a *cis, cis*-muconic acid intermediate. d = 3-dehydroshikimate dehydratase, e = protochatechuate decarboxylase, f = catechol-1,2-dioxygenase.

Figure 10 is a graphical representation showing adipic acid biosynthesis by petroselinic acid ozonolysis.

Figure 11 is a graphical representation showing 2-phenylethanol biosynthesis.

Figure 12 is a graphical representation showing the basic steps of C4 carbon assimilation.

Figure 13 is a graphical representation depicting the detection of pHBA and vanillate by HPLC in acid hydrolysed leaf samples taken from sugarcane leaves expressing the HCHL transgene. A: Indicated peaks show the presence of pHBA and vanillate in the leaf extract. Insets show thecharacteristic spectrum profiles for the respective compounds. B: Peaks produced by pHBA and vanillate synthetic standards.

Figure 14 is a graphical representation of *p*-hydroxybenzoic acid (pHBA) synthesis *in planta* can be accomplished by the introduction of *E. coli* chorismate pyruvate-lyase (CPL) or *P. fluorescens* 4-hydroxycinnamoyl-CoA hydratase/lyase (HCHL). CPL converts plastidal chorismate into pHBA whilst HCHL converts cytosolic 4-coumaroyl-CoA into 4-hydroxybenzaldehyde via a â-hydroxy thioester intermediate which is subsequently oxidized to pHBA by endogenous NAD+-linked dehydrogenases. In both instances, the majority of the resultant free acid is glucosylated and transported into vacuoles. Abbreviations: E4P, erythrose-4-phosphate; PEP, phosphoenolpyruvate; UDP-GT, UDP-glucosyltransferase.

Figure 15 is a graphical representation showing the distribution pattern of p-hydroxybenozic acid (pHBA) in UH1 at 20 weeks. (a) Leaf and internode pHBA levels are compared. There is generally more pHBA in the leaf than the stalk and the content in older tissue is generally higher than younger tissue. (b) pHBA levels at specific locations in the tissue are compared. The largest quantities of pHBA are found in the leaf lamina and the rind tissue of the stem. (LL = leaf lamina; LM =leaf midrib; R = rind; P = pith; VB = vascular bundles of stem tissue).

Figure 16 is a photographic representation showing a comparison of the growth phenotype between the highest pHBA producer, UH98 and the control line TC1 reveals no obvious differences. (a) Plants of approximately equivalent age were compared. TC1, (left),UH98, (right). Inset: A close-up view of the under surface of a leaf is shown for TC1 (b) and UH98 (c).

A summary of sequence identifiers used throughout the subject specification is provided below in Table 1:

TABLE 1

SUMMARY OF SEQUENCE IDENTIFIERS

SEQUENCE ID NO:	DESCRIPTION
1	Nucleotide sequence (*pha*A) encoding PhaA without signal sequence (hence, the PhaA remains in the cytosol)
2	Amino acid sequence of PhaA without signal sequence
3	Nucleotide sequence (*pha*A) encoding PhaA without signal sequence (hence, the PhaA remains in the cytosol) modified at 5' and 3' ends for insertion into a vector
4	Nucleotide sequence (*pha*B) encoding PhaB without signal sequence (hence, the PhaB remains in the cytosol)
5	Amino acid sequence of PhaB without signal sequence
6	Nucleotide sequence (*pha*B) encoding PhaB without signal sequence (hence, the PhaB remains in the cytosol) modified at 5' and 3' ends for insertion into a vector
7	Nucleotide sequence (*pha*C) encoding PhaC without signal sequence (hence, the PhaC remains in the cytosol)
8	Amino acid sequence of PhaC without signal sequence
9	Nucleotide sequence (*pha*C) encoding PhaC without signal sequence (hence, the PhaC remains in the cytosol) modified at 5' and 3' ends for insertion into a vector
10	Nucleotide sequence (*pha*A) encoding PhaA targeted to plastid
11	Amino acid sequence of PhaA with signal sequence to target to the plastid
12	Nucleotide sequence (*pha*A) encoding PhaA targeted to plastid modified at 5' and 3' ends for insertion into a vector
13	Nucleotide sequence (*pha*B) encoding PhaB targeted to plastid
14	Amino acid sequence of PhaB with signal sequence to target to the plastid
15	Nucleotide sequence (*pha*B) encoding PhaB targeted to plastid modified at 5' and 3' ends for insertion into a vector
16	Nucleotide sequence (*pha*C) encoding PhaC targeted to plastid

SEQUENCE ID NO:	DESCRIPTION
17	Amino acid of PhaC with signal sequence to target to the plastid
18	Nucleotide sequence (*pha*C) encoding PhaC targeted to modified at 5' and 3' ends for insertion into a vector
19	Nucleotide sequence (*pha*C1) encoding PhaC1 without signal sequence (hence, the PhaC1 remains in the cytosol)
20	Amino acid sequence of PhaC1 without signal sequence
21	Nucleotide sequence (*pha*C1) encoding PhaC1 without signal sequence modified at 5' and 3' ends for insertion into a vector
22	Nucleotide sequence (*pha*C1) encoding PhaC1 targeted to the peroxisome
23	Amino acid sequence of PhaC1 with signal sequence to target to the peroxisome
24	Nucleotide sequence (*pha*C1) encoding PhaC1 targeted to the peroxisome modified to 5' and 3' ends for insertion into a vector
25	Nucleotide sequence of (*pha*C1) encoding PhaC1 targeted to the plastid
26	Amino acid sequence of PhaC1 with signal sequence to target to the plastid
27	Nucleotide sequence (*pha*C1)) encoding PhaC1 targeted to the plastid modified at 5' and 3' ends for insertion into a vector
28	Nucleotide sequence (*pha*G) encoding PhaG targeted to the plastid
29	Amino acid sequence of PhaG with signal sequence to target to the plastid
30	Nucleotide sequence (*pha*G) encoding PhaG targeted to the plastid modified to 5' and 3' ends for insertion into a vector
31	Nucleotide sequence (*pha*J) encoding PhaJ targeted to the peroxisome
32	Amino acid sequence of PhaJ with signal sequence to target to the peroxisome
33	Nucleotide sequence (*pha*J) encoding PhaJ targeted to the peroxisome modified to 5' and 3' ends for insertion into a vector
34	TphaF
35	PhaF
36	PhbF
37	PhcF

SEQUENCE ID NO:	DESCRIPTION
38	PhaR
39	PhbR
40	PhcR
41	PhaC1Cf
42	PhaC1Cr
43	PhaC1PF
44	PhaJF
45	PhaJR
46	PhaGF
47	PhaGR
48	SSP-F
49	SSP-R
50	primer 3
51	primer 4
52	primer 5
53	primer 6

A list of abbreviations used herein is provided in table 2.

TABLE 2

Abbreviations

ABBREVIATION	Description
1,3-PD	1,3-Propanediol
2-PE	2-Phenylethanol
CAM	Crassulacean Acid Metabolism
DW	Dry weight
GFOR	Glucose-Fructose Oxidoreductase
GFP	Green Fluorescent Protein
PEP	Phosphoenolpyruvate
PEPcase	Phosphoenolpyruvate carboxylase
PET	Polyethylene terephthalate
sPET	Isosorbide PET
PHA	Polyhydroxyalkanoate
PHB	Polyhydroxybutyrate
pHBA	poly hydroxybenzaldehyde
PHV	Polyhydroxyvalerate / Poly-3-hydroxypentanoate
Rubisco	Ribulose bisphosphate carboxylase/oxygenase
WT	Wild type

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention provides a plant-derived bioreactor system, which is capable of withstanding the extra metabolic load placed on individual plants without resulting in deleterious growth effects. Hence, until the advent of the present invention, plants have been engineered to make small quantities of a desired product, their capacity to do so being limited, and their ability to go on doing so prevented, by these deleterious growth effects. In accordance with the present invention, certain plants were selected on the basis of the presence of highly efficient photosynthetic mechanisms for the assimilation of carbon, particular metabolic reserves, and/or useful metabolic and/or biosynthetic pathways. Grasses, and particularly C4 grasses, were identified as having particularly efficient mechanisms for the assimilation of carbon, a high growth rate and high accumulation of biomass and hence are useful as bioreactors for the production of a wide range of products. Furthermore, the C4 grass, sugarcane, is particularly useful, as this plant stores sugars in dimeric and/or polymeric forms that may be utilized when needed, for example, to supply energy during times of significant environmental stress. This store of carbohydrate would provide a ready supply of substrate for many metabolic pathways, and utilization of this store would not stress the producing plant. Therefore, the present invention is predicated, in part, on the identification of a subset of plants, namely the C4 grasses, and particularly sugarcane, as useful bioreactors on the basis of their high carbon assimilation rate, rapid growth, high biomass production and large store of carbohydrate.

Accordingly, one aspect of the present invention provides a method for generating a plant-based bioreactor system, said method comprising selecting a plant having a high efficiency carbon assimilation mechanism, rapid growth rate and/or high biomass production and/or reserves of metabolites or having a capacity to generate such reserves and/or which possess metabolic and/or biosynthetic pathways useful in the manufacture of a product of interest or a precursor form thereof; genetically modifying cells of the plant to enable access to said metabolites and/or metabolic or biosynthetic pathways; and then regenerating a genetically modified plant from said cells.

The terms "grass", "grasses" and the like are to be understood as reference to any member of the Gramineae plant family whether currently known or not. This family currently encompasses approximately 660 plant genera including 10,000 plant species. It will be readily apparent to the skilled artisan when examining a given plant, either known or novel, whether the plant is a grass.

Preferably, the grass is a C4 grass.

To minimise losses of carbon and nitrogen resulting from photorespiration, some plants such as corn and sugarcane that grow in hot climates have a different system for fixing CO_2, called C4 photosynthesis, than plants that grow in more temperate climates (which have C3 photosynthesis). The leaf anatomy of plants such as corn and sugarcane is different from that of temperate plants. The vascular bundles of these leaves are surrounded by a wreath of thick-walled parenchyma cells called bundle sheath cells, where most of the carbon-fixation takes place.

During C4 photosynthesis, CO_2 in the mesophyll cells is condensed with a 3C compound called phosphoenolpyruvic acid (PEP), by the action of the enzyme PEP carboxylase. This produces the 4C compound oxaloacetic acid which is then converted to malic or aspartic acid. The malic or aspartic acid is then moved through plasmodesmata (at the expense of ATP) into the bundle sheath cells.

In the bundle sheath cells, the 4C compounds are decarboxylated to release CO_2 and PEP. The CO_2 collected in the many mesophyll cells is concentrated into a few bundle sheath cells. Therefore, the plants can maintain a higher concentration of CO_2 in the bundle sheath cells (where the Calvin-Benson cycle of photosythesis occurs) than it can elsewhere in the leaf. This higher concentration of CO_2 minimises photorespiration.

The C4 pathway is more expensive energetically than C3 photosynthesis, but this is offset by the resulting decrease in photorespiration (where under certain conditions, plants may lose 30% of fixed carbon,). For this reason, C4 plants are well adapted to environments

that promote high levels of photorespiration (viz. subtropical and tropical climes).

The fundamentals of C4 photosynthesis are shown schematically in Figure *. The photosynthesis processes of C4 plants are divided between mesophyll and bundle sheath cells. Two steps of C4 photosynthesis which occur in the mesophyll cells are the light-dependent reactions and a preliminary fixation of CO_2 into malate or aspartate (a 4C compound). This C4 compound is transported to the bundle sheath cells, and is decarboxylated to form CO_2 and PEP. The released CO_2 is re-fixed by Rubisco and the Calvin-Benson cycle. The PEP is then recycled back to the mesophyll cells, and the photosynthates are distributed throughout the plant.

One defining aspect of a C4 plant is "Kranz anatomy" (German for "wreath"). This term refers to the characteristic one, or two concentric, layer(s) (wreath[s]) of bundle sheath cells, around the vascular bundles of the leaves. The mesophyll cells surrounding the bundle sheath cells fix CO_2 via PEP carboxylase to form 4C organic acids. These are transported to the the bundle sheath cells and are decarboxylated to regenerate CO_2, which is then refixed by the typical C3 photosynthetic pathway found in non-C4 plants. The system thus acts as a CO_2 pump, increasing the CO_2 concentration in the bundle sheath to a level where photorespiration is minimised.

Accordingly, for the purposes of the present invention C4 plants are to be understood as plants which exhibit at least one of the following characteristics in at least some part of the plant:

(i) Kranz anatomy;

(ii) fixation of CO_2 into a 4 carbon compound; and/or

(iii) decarboxylation of a 4 carbon compound.

For the purposes of the present invention, a given plant need not exhibit all of the above characteristics to be considered a C4 plant. For example, *Borszczowia aralocaspica* (Chenopodiaceae) has the photosynthetic features of C4 plants, yet lacks Kranz anatomy. This species accomplishes C4 photosynthesis through spatial compartmentation of photosynthetic enzymes, and by separation of two types of chloroplasts and other organelles in distinct positions within the chlorenchyma cell cytoplasm. Accordingly, insofar as the present invention relates to C4 plants, this is to be understood as those which exhibit 1, 2 or all 3 of the above characteristics.

It is also to be understood that the subject plant need not exhibit one or all of these criteria at all times. Some plant species, such as the amphibious leafless sedge, *Eleocharis vivipara*, can exhibit C3 and C4 characteristics such as those shown above depending on whether it is grown in a terrestrial or aquatic environment. In addition, the Cassava plant has photosynthetic mechanisms which are typical of a C3 plant, yet some studies have shown that both C3 and C4 enzymatic systems function in Cassava. The dominant photosynthetic pathway varies between C3 and C4 depending on temperature: at lower temperatures, photosynthesis follows a C3 path, and at higher temperatures, a C4 path.

Therefore, for the purposes of the present invention, C4 plants are to be understood as those plants capable of exhibiting one or more of the above characteristics under any given environmental condition. The present is not limited by the method of photosynthesis used at a given time by a given plant, the plant need only be capable of expressing at least one of the above characteristics associated with C4 photosynthesis.

In addition, plants utilising the Crassulacean Acid Metabolism (CAM) pathway are also to be considered within the definition of a C4 plant for the purposes of the present invention.

The term "Crassulacean" refers to the Stonecrop family (Crassulaceae) and related succulents in which this process is common. To date, plants in more than 18 different families including Cactaceae (Cactus family) and Bromeliaceae (Pineapple family) have

been shown to carry out CAM metabolism. The term "Acid" is derived from the observation that these plants accumulate large amounts of C4 organic acids in the dark.

Plants with CAM metabolism are typically adapted to dry, hot, high-light environments. CAM is largely a mechanism to conserve water. Plants in dry environments utilise CAM as they cannot afford to lose water by opening their stomata during the day. CAM plants circumvent water loss during the day by opening up the stomates at night to obtain carbon dioxide.

Carbon dioxide is accumulated in CAM plants using PEP carboxylase, and the fixed carbon is stored as 4-carbon compounds such as malate, as in C4 plants.

The CO_2 obtained during the night is stored as a C4 acid until ATP and NADPH are available the following day as a result of the light reactions of photosynthesis. The C4 acid is then decarboxylated and the CO_2 fixed by the Calvin-Benson cycle. Thus, in CAM plants there is a temporal separation of initial carbon fixation and the Calvin-Benson cycle, whereas in other C4 plants there is a spatial separation.

In summary the sequence of events in CAM plants is:

Night → stomates open → nocturnal transpiration (lower than diurnal) and carbon fixation by PEPcase → OAA produced → reduced with NADPH to malate → shuttled into vacuole as malic acid → malic acid content of vacuole increases → starch depleted to provide PEP for carboxylation → day → stomates close → transpiration decreased → malic acid content decreases → resulting malate decarboxylated to provide carbon dioxide for Calvin cycle → starch content increases.

Accordingly, as CAM plants exhibit fixation of CO_2 into a 4C compound, and decarboxylation of a 4C compound, they are to be understood as within the definition of C4 plants for the purposes of the present invention.

The present invention is particularly directed to C4 grasses; however, other non-grass C4 plants such as woody or herbaceous plants which utilise the C4 pathway are also contemplated by, and are within the scope of, the present invention.

In a preferred embodiment of the present invention the subject plant is a grass, more preferrably a C4 grass. In a particularly preferred embodiment the C4 grass is a member of the *Saccharum* genus, and particularly the *Saccharum* hybrid, sugarcane.

Commercially grown sugarcane varieties are mainly interspecific hybrids and are vegetatively propagated. There are about six different species contributing to the gene pool: *Saccharum officinarum, Saccharum robustum, Saccharum barberi, Saccharum spontaneum, Saccharum sinense* and *Erianthus sp*. Hence, the scope of present invention is not to be limited to any one variety but should be regarded as extending to and encompassing other species of *Saccharum*.

The preferred compounds to be produced by the plant bioreactor include: vanillin, sorbitol, PHAs, indigo, fructan, lactic acid, adipic acid, 1,3-propanediol and 2-phenylethanol. However, the present invention extends to the use of C4 grasses as bioreactors to generate a compounds such as therapeutics, nutrapharmaceuticals, diagnostic agents including single chain antibodies, industrial enzymes and the like. However, the present invention is in no way limited by the exemplified compounds and methods.

The present invention contemplates, therefore, a method for producing a product of interest including a product or intermediate of a biosynthetic or metabolic pathway in a C4 grass, said method comprising expressing one or more genetic sequences which encode one or more enzymes or proteins required for the production of the product or intermediate or a homolog or precursor thereof or which induces gene silencing of genetic material which encodes an enzyme or protein in a biosynthetic or metabolic pathway in cell of a C4 grass plant such that the product or intermediate accumulates in the cytosol, storage vacuole, plastid or non-plastid organelles of the cell, or accumulates in the juice or vascular fluid of the plant.

For the purposes of the present invention, the application of a plant as a bioreactor, is to be also understood as the alteration of existing plant metabolism or the introduction of new plant metabolism to generate a non-endogenous plant product or an endogenous plant product at non-native levels.

In the case of the metabolic engineering of native plant biochemical pathways, this may be achieved *via* a number of means. Alterations to the metabolic activity of an organism can be made at the gene, gene expression and protein levels.

Metabolic engineering may be affected at the protein level in an organism by the administration of particular enzyme activators or inhibitors. For example, the activity of particular biosynthetic enzymes may be regulated by the administration of particular enzyme inhibitors to the plant. These inhibitors may directly effect the accumulation of a product by reducing the activity of a particular biosynthetic enzyme. In addition indirect effects such as the redirection of metabolic flux into other pathways may be a product of a particular enzyme inhibition. For example, the blocking of a particular enzyme may lead to the buildup of an intermediate which may then be directed into a different metabolic pathway, leading to the increased accumulation of the product of the second pathway.

The present invention therefore contemplates the production of a product in a C4 grass wherein product accumulation is at least in part predicated on the direct activation or inhibition of an enzyme in a biosynthetic pathway by the administration to the plant of an enzyme inhibitor or activator. An enzyme "inhibitor" or "activator" also includes genetic inhibitors or activators; ie. nucleic acid molecules or RNAi-type molecules which induce gene silencing or a structural gene or a regulatory gene which positiviely or negatively regulates structural gene expression.

In a preferred embodiment, increased accumulation of a product or intermediate from a biosynthetic or metabolic pathway is a result of inhibition of one or more biosynthetic enzymes in the pathway or re-direction of metabolite flow down another pathway.

Particular biosynthetic or metabolic pathways may be induced in plants, parts of plants and/or plant cells in culture by the addition of elicitors, or by changes in environmental conditions. Gene expression in plants and other organisms is mediated by a number of physical, chemical and biotic factors. Physical factors such as light intensity and photoperiod have been implicated in the expression of many plant genes including genes involved in morphogenesis and plant secondary metabolism. For example, the anthocyanin biosynthetic genes, *PAL* and *CHS* are induced by increased light intensity and increased photoperiod. In a similar way, temperature and osmotic stresses have also been shown to alter gene expression in a broad range of biological systems. For example, the *KIN1*, *COR15a*, and *LTI78* genes in *Arabidopsis thaliana* are sensitive to induction by low temperatures (Knight *et al*., *Plant Cell* 11(5): 875-886, 1999), and a range of heat shock proteins and glycolytic enzymes are induced in the bacterium *Lactobacillus rhamnosus* in response to heat and osmotic stress (Prasad *et al*., *Appl. Environ. Microbiol.* 69(2): 917-925, 2003). Many of these gene expression changes are a result of stress to the cells. In addition, other physical factors such as wounding and drought have also been associated with altered gene expression in plants. For example in the fig tree *Ficus carica*, drought stress induced genes encoding a peroxidase, a chitinase and a trypsin inhibitor (Kim *et al*., *Plant Cell Physiol.* 44(4): 412-414).

Chemical inducers of gene expression have been identified for many biological systems and specific gene promoters. Examples of these include the induction of chalcone synthase, a phenylpropanoid pathway enzyme, by chemical elictors such as jasmonate. Also, bacterially synthesized lipochitooligosaccharides (Nod factors) induce the early nodulin (*ENOD*) genes in leguminous plants (Fang and Hirsch, *Plant Physiol.* 115:53-68, 1998). Several compounds have also been demonstrated to induce gene expression associated with plant defence, such as silicon dioxide, phosphate salts, and polyunsaturated fatty acids (Sticher *et al*., *Annu. Rev. Phytopathol.* 35: 235-270, 1997).

A number of biological agents have also been demonstrated to alter gene expression in other organisms. For example, many phytopathogenic fungi and bacteria induce the

expression of a number of defence-related genes, such as the PR proteins, β-glucanases, terpenoid biosynthetic enzymes and genes in the 'salicylic acid' defence pathway. Non-pathogenic, microbial colonists of plant induce yet another different set of genes in the plant host (Han *et al.*, *Phytopathology* 90: 327-332, 2000).

Alterations or changes to cultural practices, culture conditions and growth conditions, including photoperiod and/or temperature, are considered to be conditions which are not standard conditions for the growth of the plant or cell. For example, growth of a plant or plant cell culture under a 24 hour light, would be considered an altered photoperiod for the purposes of the present invention. Second, growth of a plant at a temperature substantially above or below the optimum growth temperature of said plant, plant cell culture or bacterial culture would be considered an altered temperature for the purposes of the present invention. The preceeding examples in no way limit the invention and it will be clear to those of skill in the art what constitutes altered, changed or abnormal conditions pertaining to a given cell, cell culture, plant or organism.

Methods of altering gene expression in sugarcane contemplated by the present invention are to be understood as physical processes or conditions, chemical compounds and biological, including genetic agents. Non-limiting examples of physical agents include alterations to culture and/or growth conditions of the cell or organism, light intensity and/or photoperiod, temperature, growing season, and or physical wounding. Examples of chemical agents that may alter gene expression in plants include phytohormones such as auxins, cytokinins and gibberellins; signalling molecules such as flavanoids, saccharides, sterols and peptides; herbicides and antibiotics. Examples of biological agents contemplated by the present invention include microorganisms, such as bacteria and fungi; viruses; transposons and plasmids as well as RNAi-inducing genetic molecules including a hairpin loop or other means to induce gene silencing (eg. post-transcriptional gene silencing). The preceeding examples are only illustrative in nature and in no way limit the invention to the said agents.

Accordingly, the present invention contemplates the use of sugarcane as a bioreactor.

Therefore, alteration to the gene expression profile of sugarcane to effect the production of an endogenous metabolite at an increased level, or to produce any heterologous metabolite is within the scope of the present invention. Accordingly, induction or supression of any biosythetic or metabolic genes in sugarcane, such as those exemplified herein, is to be considered within the scope of the present invention. Reference to a biosynthetic or metabolic gene also includes a regulatory gene.

The application of a plant as a bioreactor may be affected by the introduction of a new biosynthetic or metabolic pathway into the plant, or the reirection of metabolic flux down a particular pathway in a plant. For the purposes of the present invention, "introduction of a new biosynthetic or metabolic pathway" encompasses where the introduced pathway is a single protein or enzyme, which may in itself be the end-product. For example, the introduction of a nucleic acid molecule, whether or not it encodes a protein of interest, would fall within the scope of the subject invention. In this regard, a "protein" includes a protein, polypeptide or peptide as well as a glycoprotein, phosphoprotein or phospholipoprotein. Alternatively, the invention also contemplates the introduction of one or more enzymes or proteins, wherein the introduced enzyme or protein catalyses one or more reactions in the synthesis of the product of interest. For example, *inplanta* synthesis of vanillin could be introduced to sugarcane by the introduction of the enzymes feruloyl-CoA synthetase and enoyl-CoA hydratase.

Typically, the production of one or more metabolites or heterologous proteins, polypeptides or peptides in a plant is achieved by expression of a nucleic acid molecule encoding the protein, polypeptide or peptide of interest. Any nucleic acid which encodes a protein, polypeptide or peptide of interest is contemplated by the present invention. However, preferred nucleic acids include those encoding:

(i) vanillin biosynthetic enzymes, including 3-dehydroshikimate dehydratase, catechol-o-methyltransferase, aryl aldehyde dehdrogenase, feruloyl-CoA synthetase, enoyl-CoA hydratase/aldolase;

(ii) sorbitol biosynthetic enzymes, including glucose/fructose oxidoreductase;

(iii) PHA biosynthetic enzymes, including 3-ketothiolase, acetoacetyl-CoA reductase, PHA synthase, enoyl hydratase, 3-hydroxyacyl-acyl carrier protein:CoA tranferase;

(iv) indigo biosynthetic enzymes, including tryptophanase, L-tryptophan indole lyase, napthalene dioxygenase, *R. eutrophica bec* gene product;

(v) fructan biosynthetic enzymes, including fructosyltransferases and levansucrases;

(vi) lactic acid biosynthetic enzymes, including lactate dehydrogenase;

(vii) adipic acid biosynthetic enzymes, including 3-dehydroshikimate dehydratase, protocatechuate decarboxylase and catechol 1,2-dioxygenase;

(viii) petroselinic acid biosynthetic enzymes, including 3-ketoacyl-ACP synthase;

(ix) 1,3-propanediol biosynthetic enzymes including glycerol dehydratase, 1,3-propanediol oxidoreductase, glycerol-3-phosphate dehydrogenase, and glycerol-3-phosphatase; and/or

(x) 2-phenylethanol biosynthetic enzymes including aromatic-L-amino acid decarboxylase, 2-phenylethylamine oxidase and aryl alcohol dehydrogenase.

(xi) pHBA biosynthetic enzymes including 4-hydroxycinnamoyl hydratase/lyase and chorismate pyruvate lyase.

Nucleic acids encoding a particular protein or enzyme may be chemically synthesised or

isolated from another organism. In a preferred embodiment of the present invention, the gene encoding a protein or enzyme of interest is isolated from bacteria, fungi, animals, plants, protists or archaea. Bacteria provide a convenient source of genes encoding useful enzymes and proteins, although the present invention should not be limited by the source of the gene encoding the protein or enzyme of interest. Particularly useful microorganisms for the isolation of useful genes in the context of the present invention include: *R. eutropha, Aeromonas spp., Pseudomonas aeruginosa, Rhodococcus ruber, Nocardia corallina, Zymomonas mobilis, Enterobacter aerogenes, Pseudomonas putida, Bacillus subtilis, Klebsiella pneumoniae, Acinetobacter calcoaceticus* the actinomycetes (particularly *Streptomyces* spp.), *Escherichia coli* and yeast such as *Saccharomyces cereviseae*. However, it should be noted that these microorganisms represent only examples of the possible source of a gene encoding a protein or enzyme of interest, and the present invention is in no way limited by the source of the nucleic acid encoding the protein or enzyme of interest. Furthermore, as indicated above, the nucleic acid molecule may not necessarily encode an enzyme or protein but may encode a sense RNA or an antisense RNA, for use in gene silencing for example.

In order to maximise transcription, and/or transcript stability and/or translation and/or post-translational stability of the gene productof a heterologous gene in a plant, particularly a gene from a bacterium, it may be necessary to alter the sequence of the gene. For example, to maximise translation of the gene transcript, it may be necessary to alter the coding sequence of the gene to reflect the preferred codon usage of the host plant. Similarly, to maximise translation of the gene transcript it may be necessary to alter the sequence context of the gene's translation initiation site to reflect the preferred sequence context recognised by the host's translational machinery. Similarly, it may be necessary to add 5' and/or 3' non-translated regions to the coding sequence of the gene to maximise transcript stability within the host cell. Similarly, it may be necessary to alter the coding sequence of the gene to remove cryptic protease-recognition sites. Therefore, the present invention encompasses genes isolated from a bacterial, fungal, animal, protist or archaeal source, which have undergone modification to maximise transcription, and/or transcript stability, and/or translation, and or post-translational stability in a plant host. Other methods for the

alteration of the sequence of a gene will be readily ascertained by those of skill in the art and need not be elaborated further here.

In a preferred embodiment of the present invention, a number of products may be produced according to any of a number of methods including those herein described. Examples of compounds that may be produced *via* metabolic engineering of a subject plant include: vanillin (4-hydroxy-3-methoxybenzaldehyde); sorbitol; PHAs; indigo; fructan; lactic acid (2-hydroxypropanoic Acid); adipic acid; 1,3 propanediol and 2-phenylethanol. These compounds, however, are only examplary, and the present invention is predicated on the use of C4 plants as bioreactors for any compound that can be synthesised in the plant. Accordingly, the present invention is not limited to any one product or method for producing the product.

In a preferred embodiment of the present invention, the plant selected is a C4 grass and the product of interest is a mixture of different chain length polymers of hydroxyalkanoic acid monomers. The present invention extends, however, to the use of C4 grasses as bioreactors to generate a range of compounds such as therapeutics, nutrapharmaceuticals and diagnostic agents such as single chain antibodies.

Accordingly, the present invention contemplates a method for accumulating polymers comprising one or more species of hydroxyalkanoic acid monomer in a C4 grass, said method comprising expressing one or more genetic sequences which encode enzymes required for the production of the polymers or a homolog or precursor thereof in a cell of a C4 grass such that PHA polymers accumulate in the cytosol, storage vacuole, peroxisome and ontologically related organelles, or plastid or non-plastid organelles of said cell.

Polyhydroxyalkanoates (PHAs) are polyesters of one or more species of hydroxyalkanoic acid monomers. PHAs, which are bacterial carbon-storage polymers analogous to starch in plants and glycogen in animals, are a diverse class of compounds, with over 100 different hydroxyalkanoic acid sidechains identified to date. For example, PHB is a polymer of 3-hydroxybutyrate, and PHV is a polymer of hydroxyvalerate. PHAs can also be co-

polymers, for example, poly-(3-hydroxybutyrate-co-3-hydroxyvalerate). As defined herein, therefore, "polymers comprising one or more species of hydroxyalkanoic acid monomer" and "PHAs" are synonymous and encompass any such carbon-storage polymer.

PHAs are polymers that share many properties with petrochemicallyderived, synthetic polymers. The main advantages of PHAs over synthetic polymiers are that they are readily biodegradable and are made from renewable resources such as sugars and fatty acids.

While different bacterial species have developed different mechanisms for PHA biosynthesis, 3-hydroxyacylCoA is the precursor for them all. Figure 1 outlines four different strategies utilised by a wide range of bacteria including, for example, *R. eutropha*, *Aeromonas spp.*, *Pseudomonas aeruginosa*, *Rhodococcus ruber* and *Nocardia corallina*, to synthesize their requirements of polymers of 3-hydroxy acids as an energy source. Precursors may be generated by sugar glycolysis (strategy I), from metabolic intermediates of the β-oxidative (strategy II) and biosynthetic fatty acid (strategy III) pathways and from metabolites in other pathways such as the methylmalonyl-CoA pathway (strategy IV).

In the bacterium *R. eutropha*, for example, strategy I is used and the PHB biosynthetic pathway consists of three steps catalyzed by three enzymes, respectively: first, two molecules of acetyl CoA are condensed to acetoacetyl-CoA by 3-ketothiolase (encoded by *pha*A); secondly, aceotacetyl-CoA is reduced to D-3-hydroxybutyryl-CoA by acetoacetyl-CoA reductase (encoded by *pha*B); and thirdly, D-3-hydroxybutyryl-CoA is polymerized into PHB by PHB synthase (encoded by *pha*C). The genes are clustered in a single operon in the order *pha-CAB* A diagrammatic representation of this particular biosynthetic pathway is provided in Figure 2.

Aeromonas spp. are examples of organisms that employ strategy II. They express a (R)-specific enoyl hydratase (PhaJ). This enzyme catalyzes the formation of 3-hydroxyacyl-CoA from enoyl-CoA intermediates in β-oxidation of fatty acids, thereby generating substrates for the polymerase. Alternatively, organisms like *P. aeruginosa* produce PHAs from intermediates in the *de novo* fatty acid biosynthetic pathway (strategy III). A 3-

hydroxyacyl-acyl carrier protein:CoA transferase enzyme designated PhaG converts 3-hydroxyacyl-acyl carrier protein to its CoA analog. The hydroxyacyl-CoA is then incorporated into the nascent polymer. Finally, organisms such as *R. ruber* and *N. corallina* are able to generate precursors for PHA production from the methylmalonly-CoA pathway (strategy IV).

Many other micro-organisms have developed alternative biosynthetic pathways for the manufacture of PHAs, including PHB, to meet their energy-requiring needs. Biosynthetic genes from almost 40 different organisms have been cloned. Only limited homologies are exhibited at both the nucleotide and amino acid levels, which is not surprising considering the large number of PHAs naturally produced by bacteria. The structural organisation of loci encoding PHA genes is equally diverse.

Any one or more or a combination of these systems may be adapted to provide suitable genetic sequences for use in accordance with the present invention. Consequently, genetic sequences which "encode enzymes required for the production of polymers" of hydroxyalkanoic acids as used herein in the context of the present invention may comprise a combination of one or more of any sequence wherein the enzyme or enzymes thereby encoded usually operate *in vivo* singly or together to effect the biosynthesis of PHAs, including PHB.

Preferred suitable genetic sequences comprise a combination of one or more of *pha*A, *pha*B, *pha*C, *pha*C1, *pha*G, *pha*J. These genes encode enzyme products referred to herein as PhaA, PhaB and PhaC, PhaC1, PhaG and PhaJ.

In one preferred embodiment, the C4 grass is engineered to express one or more of *pha*A, *pha*B and *pha*C such that it does not accumulate the PHAs in the plastid. In another preferred embodiment, the plant is engineered to express, in addition, one or more of *pha*C1, *pha*G and *pha*J, such that it accumulates the PHAs in the plastid.

The nucleotide sequences encoding *pha*A, *pha*B and *pha*C may come from any suitable

source but the genes from *R. eutrophia* are particularly useful in the practice of the present invention. The nucleotide sequences for these genes are given in SEQ ID NO:1 (*pha*A), SEQ ID NO:4 (*pha*B) and SEQ ID NO:7 (*pha*C) where the nucleotide sequence does not include a signal sequence and, hence, the products, i.e. PhaA, PhaB and PhaC, respectively, are located in the cytosol.

Nucleotide sequences of *pha*A, *pha*B and *pha*C with a signal sequence to direct the products to the plastid are shown in SEQ ID NO:10, SEQ ID NO:13 and SEQ ID NO:16, respectively.

The nucleotide sequences encoding *pha*C1, *pha*G and *pha*J may likewise come from any suitable source. In the case of *pha*C1, *P. aeruginosa* provides a suitable source. Nucleotide sequences encoding *pha*G and *pha*J may be derived from, for example, *Pseudomonas putida* and *Aeromonas caviae*, respectively. The nucleotide sequence of *pha*C1 is given in SEQ ID NO:19, where the nucleotide sequence does not include a signal sequence and, hence, the product, i.e. PhaC1 is located in the cytosol.

Nucleotide sequences of *pha*C1, *pha*G and *pha*J with a signal sequence to direct the products to the peroxisome and plastid (*pha*C1), and to the plastid (*pha*G) and peroxisome (*pha*J), respectively, are shown in SEQ ID NO:22 and SEQ ID NO:25 and SEQ ID NO:28 and SEQ ID NO:31, respectively.

Clearly, the genetic sequences may be modified to insert any leader or tail sequence to direct the enzyme to an appropriate location in the cell.

Another aspect of the present invention contemplates a method for producing PHAs in a C4 grass, said method comprising expressing one or more genetic sequences comprising *pha*A, *pha*B, *pha*C, *pha*C1, *pha*G and/or *pha*J or a derivative or homolog of any one of these in cells of a C4 grass such that polymers of a PHA accumulate in the cytosol, storage vacuole, plastid or non-plastid organelle of said cell.

Where accumulation is in the cytosol, the PHA is preferably a PHB.

A homolog of a *pha*A, *pha*B, *pha*C, *pha*C1, *pha*G and *pha*J includes nucleotide sequences having at least about 60% identity to one of SEQ ID NO:1, SEQ ID NO:4, SEQ ID NO:7, SEQ ID NO:19, SEQ ID NO:28 or SEQ ID NO:31 (or one of SEQ ID NO:10, SEQ ID NO:13, SEQ ID NO:16, SEQ ID NO:22 or SEQ ID NO:25) after optimal alignment or nucleotide sequences capable of hybridizing to SEQ ID NO:1, SEQ ID NO:4, SEQ ID NO:7, SEQ ID NO:19, SEQ ID NO:28, SEQ ID NO:31, SEQ ID NO:10, SEQ ID NO:13, SEQ ID NO:16, SEQ ID NO: 22 or SEQ ID NO:25 or their complementary forms under low stringency conditions.

Alternatively, or in addition, a homolog at the amino acid level includes an enzyme having an amino acid sequence with at least about 60% identity to SEQ ID NO:2, SEQ ID NO:5, SEQ ID NO:8, SEQ ID NO:11, SEQ ID NO:14, SEQ ID NO:17, SEQ ID NO:20, SEQ ID NO:23, SEQ ID NO:26, SEQ ID NO:29 or SEQ ID NO:32.

Preferably, percentage similarities include at least about 70%, at least about 80%, at least about 90% and at least about 95% or above at the nucleotide and amino acid sequence levels such as 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 and 100%.

Accordingly, reference herein to *pha*A, *pha*B, *pha*C *pha*C1, *pha*G, *pha*J or PhaA, PhaB, PhaC, PhaC1, PhaG and PhaJ includes all homologs thereof.

In another embodiment, the present invention contemplates a method for generating a plant which produces PHAs, said method comprising introducing into cells of said plant a genetic sequence comprising:-

(i) a nucleotide sequence encoding a phaA or homolog thereof;

(ii) a nucleotide sequence encoding phaB or homolog thereof;

(iii) a nucleotide sequence encoding phaC or homolog thereof;

(iv) a nucleotide sequence encoding phaC1 or homolog thereof;

(v) a nucleotide sequence encoding phaG or homolog thereof;

(vi) a nucleotide sequence encoding phaJ or homolog thereof

(vii) SEQ ID NO:1 or SEQ ID NO:3 or SEQ ID NO:10 or SEQ ID NO:12 or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to SEQ ID NO:1 or SEQ ID NO:3 or SEQ ID NO:10 or SEQ ID NO:12 or a complementary form thereof under low stringency conditions;

(viii) SEQ ID NO:4 or SEQ ID NO:6 or SEQ ID NO:13 or SEQ ID NO:15 or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to SEQ ID NO:4 or SEQ ID NO:6 or SEQ ID NO:13 or SEQ ID NO:15 or a complementary form thereof under low stringency conditions;

(ix) SEQ ID NO:7 or SEQ ID NO:9 or SEQ ID NO:16 or SEQ ID NO:18 or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to SEQ ID NO:7 or SEQ ID NO:9 or SEQ ID NO:16 or SEQ ID NO:18 or a complementary form thereof under low stringency conditions;

(x) SEQ ID NO:19 or SEQ ID NO:21 or SEQ ID NO:22 or SEQ ID NO:24 or SEQ ID NO:25 or SEQ ID NO:27 or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to SEQ ID NO:19 or SEQ ID NO:21 or SEQ ID NO:22 or SEQ ID NO:24 or

SEQ ID NO:25 or SEQ ID NO:27 or a complementary form thereof under low stringency conditions;

(xi) SEQ ID NO:28 or SEQ ID NO:30 or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to SEQ ID NO:28 or SEQ ID NO:30 or a complementary form thereof under low stringency conditions;

(xii) SEQ ID NO:31 or SEQ ID NO:33 or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to SEQ ID NO:31 or SEQ ID NO:33 or a complementary form thereof under low stringency conditions;

and then regenerating a plant from said cells.

Preferably, the plant is a C4 grass and, in a particularly preferred embodiment, a sucrose-storing monocotyledonous plant. This aspect of the present invention includes progeny of the first generation plants.

A convenient C4 grass for use in the present invention is sugarcane. Sugarcane has certain advantages which make it a useful crop for use as a bioreactor including, *inter alia*, its efficient carbon fixation, high biomass accumulation, rapid growth in subtropical and tropical climates, natural ability to accumulate large quantities of sucrose, hardiness, and ease of growth. In addition, a micropropagation system is already available and an extensive industry infrastructure exists.

In order that PHAs may be produced in cells of a C4 grass, suitable sequences such as those derived from *R. eutropha* must be introduced into and expressed in the cells. That is, the plant needs to undergo genetic modification so that the metabolites and/or metabolic and/or biosynthetic pathways can be harnassed for the production of the PHAs. This may conveniently be achieved through the use of genetic constructs, engineered to comprise

nucleotide sequences required to effect PHA production.

In another preferred embodiment of the present invention, the plant selected is a C4 grass and the product of interest is pHBA...

In another preferred embodiment of the present invention, the plant selected is a C4 grass and the product of interest is vanillin.

Vanillin (4-hydroxy-3-methoxybenzaldehyde) may be produced as a co-product with sucrose in sugarcane. A number of biological pathways have been discovered for the biosynthesis/biodegradation of vanillin. At least 2 of these have substrates which are available in plants, *viz*:

(i) 3-dehydroshikimic acid is a compound which is produced as an intermediate in the shikimate pathway. A pathway has been determined which converts this substrate *via* 3-dehydroshikimate dehydratase to protocatechuic acid then to vanillic acid *via* catechol-o-methyltransferase and finally to vanillin *via* aryl aldehyde dehydrogenase.

(ii) Ferulic acid is a secondary metabolite of the phenylpropanoid pathway involved in lignin synthesis. It is converted *in planta* to feruloyl-CoA by feruloyl-CoA synthetase which in turn is converted to vanillin by enoyl-CoA hydratase/aldolase.

Accordingly, the present invention further contemplates a method for producing vanillin in a C4 grass, said method comprising expressing one or more genetic sequences which encode enzymes required for the production of vanillin, or a homolog or precursor thereof in a cell of a C4 grass such that the vanillin accumulates in the cytosol, storage vacuole, plastid or non-plastid organelle, or is extra-cellularly secreted.

Either of these pathways, or a combination of these pathways, may be adapted to provide

suitable genetic sequences for use in the production of vanillin or precursors thereof in Sugarcane. Consequently, genetic sequences which "encode enzymes required for the production of vanillin" as used herein in the context of the present invention may comprise a combination of one or more of any sequence wherein the enzyme or enzymes thereby encoded usually operate *in vivo* singly or together to effect the biosynthesis of vanillin or a precursor thereof.

Clearly, the genetic sequences may be modified to insert any leader or tail sequence to direct the enzyme to an appropriate location in the cell.

Another aspect of the present invention contemplates a method for producing Vanillin in a C4 grass, said method comprising expressing one or more genetic sequences encoding 3-dehydroshikimate dehydratase, catechol-o-methyltransferase, aryl aldehyde dehydrogenase, feruloyl-CoA synthetase, enoyl-CoA hydratase and/or enoyl-CoA aldolase, in cells of a C4 grass such that vanillin accumulates anywhere in the cell or extra-cellular matrix of the plant.

Accordingly, reference herein to 3-dehydroshikimate dehydratase, catechol-o-methyltransferase, aryl aldehyde dehydrogenase, feruloyl-CoA synthetase, enoyl-CoA hydratase and/or enoyl-CoA aldolase, includes all homologs thereof.

In another embodiment, the present invention provides a method for generating a plant which produces vanillin or a precursor thereof, said method comprising introducing into cells of said plant a genetic sequence comprising at least one of the following:-

(i) a nucleotide sequence encoding a 3-dehydroshikimate dehydratase and/or;

(ii) a nucleotide sequence encoding catechol-o-methyltransferase;

(iii) a nucleotide sequence encoding aryl aldehyde dehydrogenase;

(iv) a nucleotide sequence encoding feruloyl-CoA synthetase;

(v) a nucleotide sequence encoding enoyl-CoA hydratase ;

(vi) a nucleotide sequence encoding enoyl-CoA aldolase;

(vii) a nucleotide sequence encoding a homolog of any one of (i) though (vi)

and then regenerating a plant from said cells.

Preferably, the plant is a C4 grass and, in a particularly preferred embodiment, sugarcane. Included in this aspect of the present invention are progeny of the first generation plants.

In order that Vanillin may be produced in cells of a C4 grass, suitable sequences such as those encoding one or more Vanillin biosynthetic enzymes must be introduced into and expressed in the cells. That is, the plant needs to undergo genetic modification so that the metabolites and/or metabolic and/or biosynthetic pathways can be harnessed for the production of the vanillin or a precursor thereof. This may conveniently be achieved through the use of genetic constructs, engineered to comprise nucleotide sequences required to effect vanillin production.

In another preferred embodiment of the present invention, the plant selected is a C4 grass and the product of interest is sorbitol.

Sorbitol is a polyol that is found naturally in many fruits. To satisfy the high demand for this compound it is synthesized industrially by hydrogenation of corn-derived glucose in aqueous solution using nickel-containing catalysts. The majority of the sorbitol produced is consumed in the manufacture of toothpaste, confectionary, and ascorbic acid.

A new market for sorbitol in the polymer sector has been described by industry. Isosorbide or 1,4-3,6-dianhydrosorbitol is produced by the acid catalyzed dehydration of sorbitol.

Recent patents have demonstrated that copolymers of isosorbide and polyethylene terephthalate (PET) exhibit superior strength and rigidity compared to PET alone. 4.4 billion lb of PET is currently used in food and beverage containers (Source: US Dept. Energy, 2001). Replacing PET with the isosorbide-PET copolymer (sPET) would reduce the overall consumption of petroleum-derived PET because less sPET is needed to achieve the equivalent strength. The projected sPET production is 1 billion lb per year by 2020, utilizing 100 million lb of isosorbide (Source: US Dept. Energy, 2001).

Sorbitol is already produced from a renewable feedstock. The main incentive to use sugarcane as a sorbitol biofactory is to capitalize upon a potential future demand for this product by the plastics industry.

Sorbitol can also be converted to other useful chemicals. Propylene glycol, ethylene glycol, and glycerol can be derived from catalytic hydrogenolysis of sorbitol. These chemical feedstocks are currently derived from petrochemicals. .

The biosynthesis of sorbitol produces the coproduct gluconolactone. The enzyme glucono-δ-lactonase can convert the gluconolactone into gluconic acid. Gluconic acid is used as a food acidulant, antioxidant, and a clarifier in wines and softdrinks.

Zymomonas mobilis is able to produce sorbitol from sucrose or a mixture of glucose and fructose in a one-step reaction catalysed by the glucose-fructose oxidoreductase GFOR (Genbank accession no. Z80356, M97379). The glucose is oxidized to gluconolactone while the fructose is reduced to sorbitol.

glucose + fructose → sorbitol + gluconolactone

Without limiting the present invention to any one method or mode of action, sorbitol production in sugarcane could be achieved by using GFOR. This involves constructing an expression cassette by fusing GFOR to the maize polyubiquitin promoter and nopaline synthase terminator and introducing the cassette into sugarcane callus by biolistic

transformaton. The *Z. mobilis* GFOR is not membrane-bound and resides in the periplasm and should work equally well as a cytosolic enzyme in sugarcane.

Sorbitol production is unlikely to be toxic in sugarcane since sorbitol is found in numerous fruits (apples, pears, plums, berries, cherries). Sorbitol functions physiologically to regulate osmotic stress hence extremely high levels may be detrimental and vacuolar storage may circumvent this problem.

Accordingly, the present invention further contemplates a method for producing sorbitol in a C4 grass, said method comprising expressing one or more genetic sequences which encode enzymes required for the production of sorbitol, or a homolog or precursor thereof in a cell of a C4 grass such that the sorbitol accumulates in the cytosol, storage vacuole, plastid or non-plastid organelle, or is secreted extra-cellularly.

In addition to the glucose-fructose oxidoreductase pathway, other nucleotide sequences may encode other enzymes suitable for use in the production of sorbitol or precursors thereof in Sugarcane. Consequently, genetic sequences which "encode enzymes required for the production of sorbitol" as used herein in the context of the present invention may comprise a combination of one or more of any sequence wherein the enzyme or enzymes thereby encoded usually operate *in vivo* singly or together to effect the biosynthesis of sorbitol or a precursor thereof.

Clearly, the genetic sequences may be modified to insert any leader or tail sequence to direct the enzyme to an appropriate location in the cell.

Another aspect of the present invention contemplates a method for producing sorbitol in a C4 grass, said method comprising expressing one or more genetic sequences encoding a glucose-fructose oxidoreductase or homolog thereof, in cells of a C4 grass such that sorbitol accumulates anywhere in the cell or extra-cellular matrix of the plant.

In a preferred embodiment, the glucose-fructose oxidoreductase is that encoded by the

nucleic acid sequence set forth in Genbank accession number Z80356 or M97379, or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to these sequences under low stringency conditions.

Accordingly, reference herein to glucose-fructose oxidoreductase, includes all homologs thereof.

In another embodiment, the present invention contemplates a method for generating a plant which produces sorbitol or a precursor thereof, said method comprising introducing into cells of said plant a genetic sequence comprising a nucleotide sequence encoding a glucose-fructose oxidoreductase or homolog thereof and then regenerating a plant from said cells.

Preferably, the plant is a C4 grass and, in a particularly preferred embodiment, sugarcane.

In order that sorbitol may be produced in cells of a C4 grass, suitable sequences such as those encoding one or more sorbitol biosynthetic enzymes must be introduced into and expressed in the cells. That is, the plant needs to undergo genetic modification so that the metabolites and/or metabolic and/or biosynthetic pathways can be harnessed for the production of the sorbitol or a precursor thereof. This may conveniently be achieved through the use of genetic constructs, engineered to comprise nucleotide sequences required to effect sorbitol production.

Clearly, the genetic sequences may be modified to insert any leader, tail or signal sequences to direct the enzyme to an appropriate location in the cell.

In another preferred embodiment of the present invention, the plant selected is a C4 grass and the product of interest is indigo.

Until the end of the 19th century, the sole source of indigo was from plants, woad (*Isatis tinctoria*) and Dyer's Knotweed (*Polygonum tinctorum*) in temperate climates and

Indigofera species in the tropics. Woad was widely grown in Europe, making some regions, especially Toulouse (France) and Erfurt (Germany), very wealthy until the end of the 16th century.

Plant-based indigo was almost entirely replaced in the 20th century by synthetic indigo. Today indigo is still regarded as a high value specialty chemical used mainly as a dye in the textile industry. It is produced synthetically from naphthalene by the Heumann synthesis reaction.

The chief incentive to use sugarcane as an indigo biofactory is to provide a manufacturing route that will produce relatively inexpensive indigo from a renewable feedstock.

Indigo production by microbial fermentation has been demonstrated by expressing the genes that mediate indigo formation in *E. coli.* The pigment is derived by converting endogenous tryptophan to indole using the *Enterobacter aerogenes* tryptophanase or L-tryptophan indole lyase EC 4.1.99.1 (Genbank accession no. D14297). Subsequently the indole is converted to indigo *via* two possible reactions.

Route A: Pseudomonas putida napthalene dioxygenase (Genbank accession no. M83949)

Route B: Ralstonia eutropha bec gene (Genbank accession no. AF306552)

These pathways are graphically depicted in **Figure 8.**

Without limiting the present invention to any one method or mode of action, indigo production in sugarcane involves constructing an expression cassette by fusing the aforementioned genes to the maize polyubiquitin promoter and nopaline synthase terminator and introducing the cassette into sugarcane callus by biolistic transformaton. Tryptophan is a product of the plant shikimate pathway, which is responsible for synthesizing lignin precursors. The cloned genes may be plastid-targeted since the

shikimate pathway reactions reside in this compartment. The available metabolic flux in this pathway is expected to be significant.

Aeration of the sugarcane juice will lead to spontaneous oxidation of indoxyl to an insoluble indigo precipitate. The solid precipitate should be easy to separate from the solution by filtration or centrifugation.

Another aspect of the present invention contemplates a method for producing indigo in a C4 grass, said method comprising expressing one or more genetic sequences encoding tryptophanase, L-tryptophan indole lyase, napthalene dioxygenase, and/or the *Ralstonia eutropha bec* gene, or homolgs thereof, in cells of a C4 grass such that indigo accumulates anywhere in the cell or extracellular matrix of the plant.

In a preferred embodiment, indigo accumulates in the plastid of the plant cell.

Accordingly, reference herein to genetic sequences encoding tryptophanase, L-tryptophan indole lyase, napthalene dioxygenase, and/or the *Ralstonia eutropha bec* geneincludes all homologs thereof.

In another embodiment, the present invention contemplates a method for generating a plant which produces indigo or a precursor thereof, said method comprising introducing into cells of said plant a genetic sequence comprising at least one of the following:-

(i) a nucleotide sequence encoding genetic sequences encoding tryptophanase;

(ii) a nucleotide sequence encoding L-tryptophan indole lyase;

(iii) a nucleotide sequence encoding napthalene dioxygenase;

(iv) a nucleotide sequence comprising the *Ralstonia eutropha bec* gene;

(v) the nucleotide sequence set forth in Genbank accession number D14279, or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to Genbank D14279 under low stringency conditions.

(vi) the nucleotide sequence set forth in Genbank accession number M83949, or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to Genbank M83949 under low stringency conditions.

(vii) the nucleotide sequence set forth in Genbank accession number AF306552, or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to Genbank AF306552 under low stringency conditions.

and then regenerating a plant from said cells.

Preferably, the plant is a C4 grass and, in a particularly preferred embodiment, sugarcane.

In order that indigo may be produced in cells of a C4 grass, suitable sequences such as those encoding one or more indigo biosynthetic enzymes must be introduced into and expressed in the cells. That is, the plant needs to undergo genetic modification so that the metabolites and/or metabolic and/or biosynthetic pathways can be harnessed for the production of the indigo or a precursor thereof. This may conveniently be achieved through the use of genetic constructs, engineered to comprise nucleotide sequences required to effect indigo production.

Clearly, the genetic sequences may be modified to insert any leader or tail sequence to direct the enzyme to an appropriate location in the cell. In a preferred embodiment the leader, tail or signal sequence directs the indigo biosynthetic enzyme to be localized in the plastid.

In another preferred embodiment of the present invention, the plant selected is a C4 grass and the product of interest is a mixture of different chain length polymers of fructose monomers, such as fructans.

Fructan, or levan as it is often called, is a fructose homopolysaccharide that is linked to a terminal glucose residue. Fructans are a storage carbohydrate in some plants such as Jerusalem artichoke and chicory. Certain bacilli can also synthesize fructans.

Despite a plethora of potential applications, this polymer is not yet widely used. Some of the possible uses cited by the literature are:

(i) Low calorie sweetener. Fructans possess a sweet taste but cannot be degraded in humans.

(ii) Dietary fibre

(iii) Bulking agent

(iv) Raw material for biodegradable plastics, detergents, and adhesives

Fructans may also be an inexpensive source of fructose in the future. The food industry is rapidly adopting fructose as the preferred sweetener over sucrose. Fructose may be up to 1.8 times sweeter than sucrose. Consequently, less fructose is needed to derive the same effect. Fructose syrup is presently obtained by hydrolysis of starch to glucose followed by enzymatic isomerization of glucose to fructose. The resultant solution is an equilibrium mixture of glucose/fructose that must be further purified by ion chromatography to obtain near pure fructose. This final step purportedly adds significantly to the cost of fructose manufacture. It would be possible to avoid this step if the starting material contained only fructose. Simple hydrolysis of fructans will yield pure fructose at a reduced cost compared with using starch as the raw material.

Incentives to use sugarcane as fructan biofactory include:

(i) To create a market for this product. A demand for fructans would develop if sufficient amounts were made available. The disadvantage of the existing fructan flora is the low harvestable weight per plant.

(ii) To provide an alternative and inexpensive route for fructose production.

(iii) In subtropical and tropical climates sugarcane exhibits fast growth and very high biomass yields. The high rate of CO_2 fixation due to C4 photosynthesis should facilitate a rapid accumulation of fructans.

(iv) Vegetative propagation ensures a stable germplasm and hence predictable product yields.

Naturally occurring fructans may contain 10 to 100,000 fructose residues. Bacteria produce the larger fructans whilst those occurring in plants are smaller. The larger polymers are desirable because they are less soluble in water and consequently easier to extract. Larger fructans will not affect the osmotic pressure in the cell to the same degree as smaller molecules. Therefore it is possible to store greater quantities of fructan before the cell is affected.

Numerous bacterial fructosyltransferases or levansucrases have been characterized such as Genbank accession no. AY150365, from *Bacillus subtilis*. These enzymes catalyze the transfer of the D-fructosyl residue from sucrose to the β-2,6-linked residues of fructan.

Sucrose $\rightarrow$ fructan + glucose

Without limiting the present invention to any one method or mode of action, fructan production in sugarcane would be achieved by constructing an expression cassette

containing levansucrase, the maize polyubiquitin promoter and nopaline synthase terminator and introducing the cassette into sugarcane callus by biolistic transformaton.

Another aspect of the present invention contemplates a method for producing fructans in a C4 grass, said method comprising expressing one or more genetic sequences encoding a bacterial fructosyltransferase or levansucrase in cells of a C4 grass such that a fructan accumulates anywhere in the cell or extra-cellular matrix of the plant.

In a preferred embodiment, fructan accumulates in the apoplast or vacuole of the plant cell.

Accordingly, reference herein to genetic sequences encoding fructosyltransferases and levansucreases includes all homologs thereof.

In another embodiment, the present invention relates to a method for generating a C4 grass plant which produces a fructan or a precursor thereof, said method comprising introducing into cells of said plant a genetic sequence comprising at least one of the following:-

(i) a nucleotide sequence encoding a fructosyltransferase or homolog thereof;

(ii) a nucleotide sequence encoding a levansucrase or homolg thereof;

(iii) the nucleotide sequence set forth in Genbank accession number AY150365, or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to Genbank AY150365 under low stringency conditions.

and then regenerating a plant from said cells.

Preferably, the plant is a C4 grass and, in a particularly preferred embodiment, sugarcane.

In order that a fructan may be produced in cells of a C4 grass, suitable sequences such as those encoding one or more indigo biosynthetic enzymes must be introduced into and expressed in the cells. That is, the plant needs to undergo genetic modification so that the metabolites and/or metabolic and/or biosynthetic pathways can be harnessed for the production of the indigo or a precursor thereof. This may conveniently be achieved through the use of genetic constructs, engineered to comprise nucleotide sequences required to effect fructan production.

Clearly, the genetic sequences may be modified to insert any leader, tail or signal sequence to direct the enzyme to an appropriate location in the cell. In a preferred embodiment, to maximise fructan production in sugarcane, levansucrase is directed to the apoplast or vacuole to maximize access to substrate for conversion.

In another preferred embodiment of the present invention, the plant selected is a C4 grass and the product of interest is lactic acid (2-Hydroxypropanoic acid).

The world market for solvent replacement, biodegradable plastics and oxygenated chemicals derived from lactic acid exceeds US$ 10 billion (Argonne National Laboratory, US DOE).

Without limiting the present invention to any one method or mode of action, lactic acid production in sugarcane involves the following general steps:

(i) Obtain or clone lactate dehydrogenase (LDH) or a homolog therof:

(ii) Express gene in sugarcane (cytosol, therefore no targeting is required)

(iii) Regenerate plants and evaluate for lactate (or derivative) production

Traditionally, lactic acid purification has been a complex chemical process. However, recent advances have simplified this process and made it significantly cheaper. It is

anticipated that lactic acid can be removed from the post-crushing millstream without great difficulty or extensive modification of existing structures.

Accordingly, the present invention further contemplates a method for producing lactic acid in a C4 grass, said method comprising expressing one or more genetic sequences which encode enzymes required for the production of lactic acid, or a homolog or precursor thereof in a cell of a C4 grass such that the lactic acid accumulates in the cytosol, storage vacuole, plastid or non-plastid organelle, or is secreted extra-cellularly.

Genetic sequences which "encode enzymes required for the production of lactic acid" as used herein in the context of the present invention may comprise a combination of one or more of any sequence wherein the enzyme or enzymes thereby encoded usually operate *in vivo* singly or together to effect the biosynthesis of lactic acid or a precursor thereof.

Clearly, the genetic sequences may be modified to insert any leader or tail sequence to direct the enzyme to an appropriate location in the cell.

In a preferred embodiment the lactate dehydrogenase nucleic acid sequence is expressed without a signal sequence such that the enzyme is active in the cytosol.

Another aspect of the present invention contemplates a method for producing lactic acid in a C4 grass, said method comprising expressing one or more genetic sequences encoding lactate dehydrogenase or a homolog thereof in cells of a C4 grass such that lactic acid accumulates anywhere in the cell or extracellular matrix of the plant.

Accordingly, reference herein to lactate dehydrogenase, includes all homologs thereof.

In another embodiment, the present invention contemplates a method for generating a plant which produces lactic acid or a precursor thereof, said method comprising introducing into cells of said plant a genetic sequence encoding lactate dehydrogenase or homolog thereof, and then regenerating a plant from said cells.

Preferably, the plant is a C4 grass and, in a particularly preferred embodiment, sugarcane.

In order that lactic acid may be produced in cells of a C4 grass, suitable sequences such as those encoding one or more lactic acid biosynthetic enzymes must be introduced into and expressed in the cells. That is, the plant needs to undergo genetic modification so that the metabolites and/or metabolic and/or biosynthetic pathways can be harnessed for the production of the lactic acid or a precursor thereof. This may conveniently be achieved through the use of genetic constructs, engineered to comprise nucleotide sequences required to effect lactic acid production.

In another preferred embodiment of the present invention, the plant selected is a C4 grass and the product of interest is adipic acid.

Adipic acid is classified as a bulk chemical and is among the top fifty chemicals produced in the US. It is principally used in the production of nylon 66, polyeurethane resins and plasticizers. Nearly 90% is used to produce nylon-6,6, a synthetic polymer developed by DuPont in the 1930's. This polyamide is formed by the condensation of adipic acid and 1,6-diaminohexane.

Adipic acid is presently produced industrially by benzene-based synthetic chemistry. Catalytic hydrogenation of benzene followed by air oxidation yields a ketone/alcohol mixture (cyclohexanone/cyclohexanol) that is further oxidized with nitric acid to produce adipic acid.

Incentives to use sugarcane as an adipic acid biofactory include:

(i) Provide a renewable feedstock for adipic acid manufacture.

(ii) Eliminate the production of the toxic nitrous oxide byproduct that accompanies traditional adipic acid synthesis.

(iii) Capitalize upon the high demand for this product.

(iv) Access to low cost molasses to produce products by fermentation.

(v) Access to low operating and infrastructure costs through co-location of an extraction facility (or fermentation facility) with a sugar mill.

(vi) In subtropical and tropical climated sugarcane exhibits fast growth and very high biomass yields. This is a prerequisite for economical bulk chemical production.

(vii) Vegetative propagation ensures a stable germplasm and hence predictable product yields.

Without limiting the present invention to any one method or mode of action, adipic acid may be produced in sugarcane by one of two approaches.

I. Synthesis from cis, cis-muconic acid

Niu *et al.* (*Biotechnol. Prog.*, 18: 201-211, 2002) describe a microbiological route for the production of adipic acid using *E. coli*. Three genes were introduced into *E. coli* to produce *cis, cis*-muconic acid that was subsequently purified from the fermentation broth and converted to adipic acid by catalytic hydrogenation (step g, 10% Pt/C, H_2, 3400 kPa, 25°C). This final step has a 97% conversion efficiency.

The synthesis of *cis, cis*-muconic acid in sugarcane involves making use of the shikimate pathway. In order to use the shikimate pathway to produce *cis, cis*-muconic acid the following biosynthetic enzymes, or homologs thereof are introduced into sugarcane:

Klebsiella pneumoniae 3-dehydroshikimate dehydratase (aroZ) – enzyme d

3-dehydroshikimate → protocatechuate

Klebsiella pneumoniae protocatechuate decarboxylase (aroY) – enzyme e

Protocatechuate → catechol

Acinetobacter calcoaceticus catechol 1,2-dioxygenase (catA) – enzyme f

Catechol + O_2 → *cis, cis*-muconic acid

Introduction of these genes into sugarcane involves constructing an expression cassette by fusing the genes described above to the maize polyubiquitin promoter and nopaline synthase terminator and introducing the cassette into sugarcane callus by biolistic transformation. Catechol is probably produced in most plants, and therefore, it may be unnecessary to clone additional copies of 3-dehydroshikimate dehydratase or protocatechuate decarboxylase. Preferrably, the cloned gene(s) are plastid-targeted since the shikimate pathway reactions reside in this compartment.

The merits of using plant secondary metabolism to synthesize interesting products have often been promoted in the literature (Verpoorte and Memelink, *Curr. Opin. Biotech.* 13: 181-187, 2002). The shikimate pathway executes a central role in plant secondary metabolism. This is one of the most active pathways in plants in terms of carbon flux owing to the fact that it is the source of lignin precursors. This makes it an attractive candidate for metabolic engineering.

II. Synthesis from petroselinic acid

Bio-based adipic acid can be obtained through ozonolysis (O_3) of petroselinic acid (18:1 $\Delta^{6\ cis}$), as depicted in Figure 10. The coproduct lauric acid is also a potential source of feedstock for detergent manufacture.

The seed oil of the coriander spice plant contains 80-90% petroselinic acid. A 36 kDa putative acyl-ACP desaturase (Genbank accession no. M93115) has been identified from coriander seed extracts and the corresponding cDNA was able to confer the ability to produce petroselinic acid in tobacco callus (Cahoon *et al. Proc. Natl. Acad. Sci. USA* 89: 11184-11188, 1992).

The metabolic pathway for producing petroselinic acid is unclear, however, evidence suggests that it is formed by the desaturation of palmitoyl-ACP by the 36 kDa desaturase followed by elongation to form petroselinic acid (Cahoon and Ohlrogge, *Plant Physiol.*, 104: 827-837, 1994).

$$16{:}0\text{-ACP} \rightarrow 16{:}1\ \Delta^4\text{-ACP} \rightarrow 18{:}1\ \Delta^6\text{-ACP}$$

Recent studies have identified a 3-ketoacyl-ACP synthase (Genbank accession no. AF263992) associated with the two-carbon elongation of 16:1 Δ^4-ACP (Mekhedov *et al.*, Plant Mol. Biol. 47: 507-518, 2001).

Cis, cis-muconic acid in sugarcane juice would be converted to adipic acid by catalytic hydrogenation. The adipic acid in the resultant solution can be recovered by solvent extraction. The solution is contacted with chloroform or methylene chloride and the adipic acid recovered in the aqueous fraction. The aqueous fraction would then be evaporated to yield crystalline adipic acid.

Accordingly, the present invention further contemplates a method for producing adipic acid, or a precursor thereof such as *cis*, *cis*-muconic acid, in a C4 grass, said method comprising expressing one or more genetic sequences which encode enzymes required for the production of adipic acid and/or *cis*, *cis*-muconic acid, or a homolog or precursor

thereof in a cell of a C4 grass such that adipic acid and/or *cis*, *cis*-muconic acid accumulates in the cytosol, storage vacuole, plastid or non-plastid organelle, or is secreted extra-cellularly.

Either of the hereinbefore described pathways for the production of *cis*, *cis*-muconic acid, or a combination of these pathways, may be adapted to provide suitable genetic sequences for use in the production of vanillin or precursors thereof in Sugarcane. Consequently, genetic sequences which "encode enzymes required for the production of *cis, cis*-muconic acid" as used herein in the context of the present invention may comprise a combination of one or more of any sequence wherein the enzyme or enzymes thereby encoded usually operate *in vivo* singly or together to effect the biosynthesis of *cis, cis*-muconic acid or a precursor thereof.

Clearly, the genetic sequences may be modified to insert any leader, tail or signal sequence to direct the enzyme to an appropriate location in the cell. Preferrably the leader tail or signal sequences lead to the co-localization of the adipic acid biosynthetic enzymes with the endogenous skimimate pathway enzymes in the plant. More preferably, said enzymes are localized in the plastid.

Another aspect of the present invention contemplates a method for producing *cis, cis*-muconic acid or adipic acid in a C4 grass, said method comprising expressing one or more genetic sequences encoding 3-dehydroshikimate dehydratase, protochatechuate decarboxylase, catechol 1,2-dioxygenase and/or 3-ketoacyl-ACP synthase in cells of a C4 grass such that adipic acid or an adipic acid precursor accumulates anywhere in the cell or extracellular matrix of the plant.

Accordingly, reference herein to 3-dehydroshikimate dehydratase, protochatechuate decarboxylase, catechol 1,2-dioxygenase and/or 3-ketoacyl-ACP synthase, includes all homologs thereof.

In another embodiment, the present invention contemplates a method for generating a plant

which produces adipic acid or a precursor thereof, said method comprising introducing into cells of said plant a genetic sequence comprising at least one of the following:-

(i) a nucleotide sequence encoding a 3-dehydroshikimate dehydratase and/or;

(ii) a nucleotide sequence encoding protochatechuate decarboxylase;

(iii) a nucleotide sequence encoding catechol 1,2-dioxygenase;

(iv) a nucleotide sequence encoding 3-ketoacyl-ACP synthase; and/or

(v) a nucleotide sequence encoding a homolog of any one of (i) through (iv).

and then regenerating a plant from said cells.

In a preferred embodiments:

(i) the nucleotide sequence encoding a 3-dehydroshikimate dehydatase is the *aroZ* gene from *Klebsiella pneumoniae*, or a homolog thereof;

(ii) the nucleotide sequence encoding a protochatechuate decarboxylase is the *aroY* gene from *Klebsiella pneumoniae*, or a homolog thereof;

(iii) the nucleotide sequence encoding a 1,2-dioxygenase is the *catA* gene from *Acinetobacter calcoaceticus*, or a homolog thereof;

(iv) the nucleotide sequence encoding a 3-ketoacyl-ACP synthase is the nucleotide sequence set forth in Genbank Accession number AF263992, or a nucleotide sequence having at least 60% identity thereto after optimal alignment, or capable of hybridizing to Genbank AF263992 under low stringency conditions.

Preferably, the plant is a C4 grass and, in a particularly preferred embodiment, sugarcane.

In order that adipic acid or a precursor thereof may be produced in cells of a C4 grass, suitable sequences such as those encoding one or more adipic acid biosynthetic enzymes must be introduced into and expressed in the cells. That is, the plant needs to undergo genetic modification so that the metabolites and/or metabolic and/or biosynthetic pathways can be harnessed for the production of adipic acid or a precursor thereof. This may conveniently be achieved through the use of genetic constructs, engineered to comprise nucleotide sequences required to effect the production of adipic acid and/or precursors thereof.

In another preferred embodiment of the present invention, the plant selected is a C4 grass and the product of interest is 1,3-propanediol (1,3-PD).

1,3-PD is a bifunctional alcohol that can be used as a monomer in numerous polycondensation reactions to produce polyesters, polyurethanes, and polyethers. The high cost of chemical synthesis, reportedly US$30/kg (Biebl *et al.*, *Appl. Microbiol. Biotechnol.*, 52: 289-297, 1999) has restricted its use in the past to specialty markets such as dioxane production and the solvent market.

1,3-PD is synthesized using a process in which ethylene oxide is reacted with carbon dioxide and hydrogen. An alternative method, the Degussa process, is based upon hydrolysis of acrolein followed by catalytic hydrogenation. Both routes involve the use of petrochemical feedstock.

1,3-PD is a natural product of glycerol fermentation in a few enterobacteria and clostridia.

Incentives to use sugarcane as 1,3-PD biofactory include:

(i) Provide a renewable feedstock for 1,3-PD manufacture.

(ii) Provide a high volume, low cost source of 1,3-PD to facilitate expansion of the market.

(iii) Capitalize upon the high demand for this product.

(iv) Access to low cost molasses to produce products by fermentation.

(v) Access to low operating and infrastructure costs through co-location of an extraction facility (or fermentation facility) with a sugar mill.

(vi) In subtropical and tropical climated sugarcane exhibits fast growth and very high biomass yields. This is a prerequisite for economical bulk chemical production.

(vii) Vegetative propagation ensures a stable germplasm and hence predictable product yields.

The metabolic reactions that convert glycerol to 1,3-PD have been established from *Klebsiella pneumoniae*:

Klebsiella pneumoniae glycerol dehydratase (dhaB)

glycerol → 3-hydroxypropionaldehyde + H_2O

Klebsiella pneumoniae 1,3-propanediol oxidoreductase (dhaT)

3-hydroxypropionaldehyde + NADH → 1,3-propanediol + NAD

Sugarcane does not naturally produce glycerol therefore the reactions that convert triose phosphates to glycerol must also be engineered into sugarcane.

Saccharomyces cerevisiae glycerol-3-phosphate dehydrogenase

dihydroxyacetone phosphate + NADH → glycerol-3-phosphate + NAD

Saccharomyces cerevisiae glycerol-3-phosphatase

glycerol-3-phosphate + ADP → glycerol + ATP

Without limiting the present invention to any one method or mode of action, all four new genes are cloned into sugarcane to convert it into a 1,3-PD biofactory. These genes are assembled into an expression cassette containing the maize polyubiquitin promoter and nopaline synthase terminator. The cassette is introduced into sugarcane callus by biolistic transformation and expression will be targeted to the cytosol. The accumulation of 1,3-PD in plant tissue will be assayed from plant extracts by conventional HPLC.

1,3-PD can be recovered from sugarcane juice by extraction with cyclohexane followed by vaporization of the residual solvent. Alternatively, distillation may be employed. Use of cyclohexane is environmentally unsound and distillation is energy intensive. Consequently, a method has been patented that describes the use of ion exclusion resins to recover 1,3-PD (WO0173097 Method of recovering 1,3-propanediol from fermentation broth, Archer Daniels Midland Co., 2001).

Accordingly, the present invention further contemplates a method for producing 1,3-propanediol in a C4 grass, said method comprising expressing one or more genetic sequences which encode enzymes required for the production of 1,3-propanediol, or a homolog or precursor thereof in a cell of a C4 grass such that the 1,3-propanediol accumulates in the cytosol, storage vacuole, plastid or non-plastid organelle, or is secreted extra-cellularly.

Any of the disclosed biosynthetic steps, or a combination of these, may be adapted to

provide suitable genetic sequences for use in the production of 1,3-propanediol, or precursors thereof, in sugarcane. Consequently, genetic sequences which "encode enzymes required for the production of 1,3-propanediol" as used herein in the context of the present invention may comprise a combination of one or more of any sequence wherein the enzyme or enzymes thereby encoded usually operate *in vivo* singly or together to effect the biosynthesis of 1,3-propanediol or a precursor thereof.

Clearly, the genetic sequences may be modified to insert any leader or tail sequence to direct the enzyme to an appropriate location in the cell.

The present invention contemplates a method for producing 1,3-propanediol in a C4 grass, the method comprising expressing one or more genetic sequences encoding glycerol dehydratase, 1,3-propanediol oxidoreductase, glycerol-3-phosphate dehydrogenase and glycerol-3-phosphatase in cells of a C4 grass such that 1,3-propanediol accumulates anywhere in the cell or extra-cellular matrix of the plant.

Accordingly, reference herein to glycerol dehydratase, 1,3-propanediol oxidoreductase, glycerol-3-phosphate dehydrogenase and glycerol-3-phosphataseincludes all homologs thereof.

In another aspect, the present invention contemplates a method for generating a plant which produces 1,3-propanediol or a precursor thereof, said method comprising introducing into cells of said plant a genetic sequence comprising at least one of the following:-

(i) a nucleotide sequence encoding a glycerol dehydratase and/or;

(ii) a nucleotide sequence comprising the *dhaB* gene from *Klebsiella pneumoniae*, or a homolg thereof;

(iii) a nucleotide sequence encoding 1,3-propanediol oxidoreductase;

(iv) a nucleotide sequence comprising the *dhaT* gene from *Klebsiella pneumoniae* or homolg thereof;

(v) a nucleotide sequence encoding glycerol-3-phosphate dehydrogenase;

(vi) a nucleotide sequence encoding glycerol-3-phosphatase; and/or

(vii) a nucleotide sequence encoding a homolog of any one of (i) through (iv)

and then regenerating a plant from said cells.

Preferably, the plant is a C4 grass and, in a particularly preferred embodiment, sugarcane.

In order that 1,3-propanediol may be produced in cells of a C4 grass, suitable sequences such as those encoding one or more 1,3-propanediol biosynthetic enzymes must be introduced into and expressed in the cells. That is, the plant needs to undergo genetic modification so that the metabolites and/or metabolic and/or biosynthetic pathways can be harnessed for the production of the 1,3-propanediol or a precursor thereof. This may conveniently be achieved through the use of genetic constructs, engineered to comprise nucleotide sequences required to effect 1,3-propanediol production.

Clearly, the genetic sequences may be modified to insert any leader, tail or signal sequence to direct the enzyme to an appropriate location in the cell.

In another preferred embodiment of the present invention, the plant selected is a C4 grass and the product of interest is 2-phenylethanol (2-PE).

2-phenylethanol (2-PE) is an important flavour and fragrance compound with a rose-like odour. Most of the world's annual production of several thousand tons is synthesised by chemical means but, due to increasing demand for natural flavours, alternative production

methods are being sought (Etschmann *et al*. *Appl Micribiol Biotechnol* 59:1-8, 2002)

A biological pathway for the biosynthesis of 2-PE is presented in Figure 11.

Sugarcane has a productive phenylpropanoid pathway and should adapt readily to increased demands placed on it for synthesis of 2-PE.

Accordingly, in another aspect, the present invention further contemplates a method for producing 2-phenylethanol in a C4 grass, said method comprising expressing one or more genetic sequences which encode enzymes required for the production of 2-phenylethanol, or a homolog or precursor thereof in a cell of a C4 grass such that the 2-phenylethanol accumulates in the cytosol, storage vacuole, plastid or non-plastid organelle, or is secreted extra-cellularly.

Any of the disclosed biosynthetic steps, or a combination of these, may be adapted to provide suitable genetic sequences for use in the production of 2-phenylethanol, or precursors thereof, a C4 grass such as in sugarcane. Consequently, genetic sequences which "encode enzymes required for the production of 2-phenylethanol" as used herein in the context of the present invention may comprise a combination of one or more of any sequence wherein the enzyme or enzymes thereby encoded usually operate *in vivo* singly or together to effect the biosynthesis of 2-phenylethanol or a precursor thereof.

Another aspect of the present invention contemplates a method for producing 2-phenylethanol in a C4 grass, said method comprising expressing one or more genetic sequences encoding aromatic-L-amino acid decarboxylase, 2-phenylethylamine oxidase and aryl-alcohol dehydrogenase in cells of a C4 grass such that 2-phenylethanol accumulates anywhere in the cell or extracellular matrix of the plant.

Accordingly, reference herein to aromatic-L-amino acid decarboxylase, 2-phenylethylamine oxidase and aryl-alcohol dehydrogenase includes all homologs thereof.

In another embodiment, the present invention contemplates a method for generating a plant which produces 2-phenylethanol or a precursor thereof, said method comprising introducing into cells of said plant a genetic sequence comprising at least one of the following:-

(i) a nucleotide sequence encoding a aromatic-L-amino acid decarboxylase and/or;

(ii) a nucleotide sequence encoding 2-phenylethylamine oxidase;

(iii) a nucleotide sequence encoding aryl-alcohol dehydrogenase; and/or

(iv) a nucleotide sequence encoding a homolog of any one of (i) through (iii)

and then regenerating a plant from said cells.

Preferably, the plant is a C4 grass and, in a particularly preferred embodiment, sugarcane.

In order that 2-phenylethanol may be produced in cells of a C4 grass, suitable sequences such as those encoding one or more 2-phenylethanol biosynthetic enzymes must be introduced into and expressed in the cells. That is, the plant needs to undergo genetic modification so that the metabolites and/or metabolic and/or biosynthetic pathways can be harnessed for the production of the 2-phenylethanol or a precursor thereof. This may conveniently be achieved through the use of genetic constructs, engineered to comprise nucleotide sequences required to effect 2-phenylethanol production.

Clearly, the genetic sequences may be modified to insert any leader, tail or signal sequence to direct the enzyme to an appropriate location in the cell.

In another preferred embodiment of the present invention, the plant selected is a C4 grass and the product of interest is pHBA.

A schematic depicting the pHBA biosynthetic pathway is shown in Figure 14.

In order to effect pHBA production in sugarcane, a chloroplast-targeted version of *E. coli* situated between the maize ubi-1 promoter and *nos* terminator of the expression construct pU3z-mcs-nos, was co-bombarded with a plasmid containing a selectable marker (pUKN) into embryogenic sugarcane callus to yield the UC series of transgenic lines. The UH series of plants was generated in the same manner using an analogous expression construct that contained the ORF of the *P. fluorescens* HCHL gene. The regenerated plants were grown in a greenhouse for four weeks and were then analyzed for pHBA accumulation in leaf tissue using HPLC.

pHBA accumulated in the transformed plants as two glucose conjugates, ie, a phenolic glucoside and a glucose ester. Both compounds contained a single glucose molecule that was attached by a 1-O-□-D linkage to the hydroxyl or carboxyl group of pHBA. The predominant product in all of the plants examined was the phenolic glucoside, which accounted for at least 90% of the pHBA.

The mean value for the population was 0.41% ± 0.04% of dry weight (DW), which is almost 30-fold higher than the mean value for the non-transgenic control plants 0.014% ± 0.01% DW. More important, the pHBA glucoside content of the best plant was 1.5% DW, which is equivalent to 0.69% DW free pHBA after correcting for the attached glucose molecule. This value is three times higher than the highest value obtained with transgenic tobacco plants expressing a different chloroplast-targeted version of CPL. The HCHL-expressing sugarcane plants accumulated even higher levels of pHBA. The mean value for total pHBA glucose conjugates in the UH lines was 0.70% ± 0.07% DW, and the highest level observed at this stage of development was 2.6% DW.

Based on the results obtained with the 4-week-old plants, a subset of the primary transformants was selected for further evaluation, and leaf levels of pHBA were determined after 16 weeks additional growth. Included in this analysis were the two CPL-

expressing plants that previously exhibited the highest levels of product accumulation (UC63 and UC65) and five HCHL-expressing plants. The methanol-extracted samples were subjected to acid hydrolysis, which quantitatively hydrolyzes both pHBA glucose conjugates, and free pHBA was determined by HPLC.

It was anticipated that pHBA production would continue throughout development and that the 20-week-old plants would have higher levels of pHBA glucosides than the 4-week-old plants. However, the increase in pHBA content with age was not very dramatic nor was it universally observed when product accumulation was expressed on a dry weight basis (Fig. 3A). Part of the explanation for this is the lower water content of the older plant leaf tissue. For example, the average dry weight to wet weight ratio for the 20-week-old plants was 0.23, while the corresponding value for the 4-week-old plants was 0.15. When this phenomenon is taken into account and product accumulation is expressed on a fresh weight basis it becomes far more apparent that pHBA levels did increase as the plants continued to grow (Fig. 3B), except for the two CPL-expressing plants.

The 20-week-old primary transformants were large enough to screen for stalk levels of pHBA without damaging the plants. At this stage of development, the oldest stem tissue is semi-mature and new tillers emerge. Since the stalk is the only part of the sugarcane plant that is normally harvested in the existing sugar mill infrastructure, pHBA accumulation in this tissue is the most important gauge for technical success. Leaf and stem samples were taken from 20-week-old plants, and total pHBA was determined by HPLC after methanol extraction and acid hydrolysis. The third internode from the bottom of the plant was the source of stem tissue for this analysis, and the leaf samples were obtained from the third fully unfurled leaf from the top of the plant. Generally speaking, leaf levels of pHBA were considerably higher than stalk levels.

However, the difference was much more pronounced for the CPL-expressing plants. For example, the average stalk to leaf ratio of pHBA for the five UH lines that were examined was 0.324 ± 0.031, and the highest stalk level of pHBA was 0.24% DW, which is equivalent to 0.52% pHBA glucose conjugates. In marked contrast, the corresponding

ratios for UC63 and UC65 were 0.135 and 0.133, respectively, and product accumulation in the stalk of the best plant (UC63) was only 0.06% DW. Since there are no reported values in the literature for pHBA levels in stem tissue for transgenic plants expressing CPL or HCHL, it will be very interesting to see if these observations will extend to other plant systems. Nevertheless, taken together the above results suggest that HCHL is a better catalyst for pHBA production in sugarcane than CPL, and subsequent studies focused on the UH series of plants.

To gain a better understanding of pHBA accumulation in different parts the plant, leaf and stem segments were sampled from the primary shoot of 20-week-old UH1. The first leaf at the top with a fully visible dewlap was designated "leaf 1" and consecutive leaves down the stalk were numbered in ascending order. The stem segments were numbered similarly with "internode 1" corresponding to the internode immediately above the point of attachment of leaf 1. Note that the values shown refer to total pHBA after acid hydrolysis. Except for the youngest leaf examined, product accumulation in leaves was relatively uniform along the length of the plant achieving a maximum value of ~1.0% DW. Product accumulation also varied along the length of the leaf, with the tip of the leaf having about twice as much pHBA as the base of leaf. A similar trend was observed in the stalk, but there was a much larger discrepancy between young stem tissue and old stem tissue. In agreement with the results described above, pHBA levels in mature stem tissue were about 3-fold lower than mature leaf tissue. These results add additional support to the notion that pHBA accumulation in HCHL-expressing sugarcane plants increases as a function of time.

Additional insight on pHBA distribution was obtained from dissection experiments. Three different compartments of the stalk were examined: rind, pith, and vascular bundles. The most pHBA was found in the rind (1% DW), while the pith and vascular bundles had 3- to 4-fold lower levels. Indeed, pHBA levels in the rind were very similar to values obtained from the leaf midrib and leaf lamina.

Of all of the HCHL-expressing primary transformants monitored, UH98 consistently had the highest levels of pHBA in both leaf and stem tissue. When this plant was 20 weeks old

pHBA accumulation in leaf tissue was 2.8% DW (leaf lamina, 3.35% DW; leaf midrib, 1.61% DW). The corresponding value for mature stem tissue was 0.67% DW (rind, 0.96% DW; pith, 0.65% DW). Despite these very high levels of pHBA glucose conjugates, UH98 was morphologically indistinguishable from the non-transformed control line TC1 (Fig. 5).

The present invention contemplates a method for producing pHBA in a C4 grass, the method comprising expressing one or more genetic sequences encoding one or more pHBA biosynthetic enzymes in cells of a C4 grass such that pHBA accumulates anywhere in the cell or extra-cellular matrix of the plant.

Accordingly, reference herein to hydroxycinnamoyl-CoA hydratase/lyase or chorismate pyruvate lyase includes all homologs thereof.

In a preferred embodiment, the present invention contemplates a method for generating a plant which produces pHBA or a precursor thereof, said method comprising introducing into cells of said plant a genetic sequence comprising at least one of the following:-

(i) a nucleotide sequence encoding hydroxycinnamoyl-CoA hydratase/lyase;

(ii) a nucleotide sequence encoding chorismate pyruvate lyase ;

(iii) a nucleotide sequence comprising the *ubiC* gene from *E. coli*, or a homolg thereof; and/or

(iv) a nucleotide sequence comprising the *HCHL* gene from *Pseudomonas fluorescens* or homolg thereof;

and then regenerating a plant from said cells.

Preferably, the plant is a C4 grass and, in a particularly preferred embodiment, sugarcane.

In order that pHBA may be produced in cells of a C4 grass, suitable sequences such as those encoding one or more pHBA biosynthetic enzymes must be introduced into and expressed in the cells. That is, the plant needs to undergo genetic modification so that the metabolites and/or metabolic and/or biosynthetic pathways can be harnessed for the production of the pHBA or a precursor thereof. This may conveniently be achieved through the use of genetic constructs, engineered to comprise nucleotide sequences required to effect pHBA production.

Clearly, the genetic sequences may be modified to insert any leader, tail or signal sequence to direct the enzyme to an appropriate location in the cell. In a preferred embodiment, the pHBA biosynthetic enzymes are targetted to the plastid.

To effect expression of the nucleotide sequence of the present invention, it may conveniently be incorporated into a chimeric genetic construct comprising *inter alia* one or more of the following: a promoter sequence, a 5' non-coding region, a *cis*-regulatory region such as a functional binding site for transcriptional regulatory protein or translational regulatory protein, an upstream activator sequence, an enhancer element, a silencer element, a TATA box motif, a CCAAT box motif, an upstream open reading frame, transcriptional start site, translational start site, and/or nucleotide sequence which encodes a leader sequence, and a 3' non-translated region. Preferable the chimeric genetic construct is designed for transformation of plants as hereinafter described.

The term "5' non-coding region" is used herein in its broadest context to include all nucleotide sequences which are derived from the upstream region of an expressible gene, other than those sequences which encode amino acid residues which comprise the polypeptide product of said gene, wherein 5' non-coding region confers or activates or otherwise facilitates, at least in part, expression of the gene.

The term "gene" is used in its broadest context to include both a genomic DNA region corresponding to the gene as well as a cDNA sequence corresponding to exons or a recombinant molecule engineered to encode a functional form of a product.

As used herein, the term "*cis*-acting sequence" or "*cis*-regulatory region" or similar term shall be taken to mean any sequence of nucleotides which is derived from an expressible genetic sequence wherein the expression of the first genetic sequence is regulated, at least in part, by said sequence of nucleotides. Those skilled in the art will be aware that a *cis*-regulatory region may be capable of activating, silencing, enhancing, repressing or otherwise altering the level of expression and/or cell-type-specificity and/or developmental specificity of any structural gene sequence.

Reference herein to a "promoter" is to be taken in its broadest context and includes the transcriptional regulatory sequences of a classical genomic gene, including the TATA box which is required for accurate transcription initiation, with or without a CCAAT box sequence and additional regulatory elements (i.e. upstream activating sequences, enhancers and silencers) which alter gene expression in response to developmental and/or environmental stimuli, or in a tissue-specific or cell-type-specific manner. A promoter is usually, but not necessarily, positioned upstream or 5', of a structural gene, the expression of which it regulates. Furthermore, the regulatory elements comprising a promoter are usually positioned within 2 kilobase pairs (kb) of the start site of transcription of the gene.

In the present context, the term "promoter" is also used to describe a synthetic or fusion molecule, or derivative which confers, activates or enhances expression of a structural gene or other nucleic acid molecule, in a plant cell. Preferred promoters according to the invention may contain additional copies of one or more specific regulatory elements to further enhance expression in a cell, and/or to alter the timing of expression of a gene to which it is operably connected.

The term "operably connected" or "operably linked" in the present context means placing a gene under the regulatory control of a promoter, which then controls the transcription and optionally translation of the gene. In the construction of heterologous promoter/structural gene combinations, it is generally preferred to position the genetic sequence or promoter at a distance from the gene transcription start site that is approximately the same as the

distance between that genetic sequence or promoter and the gene it controls in its natural setting, i.e. the gene from which the genetic sequence or promoter is derived. As is known in the art, some variation in this distance can be accommodated without loss of function. Similarly, the preferred positioning of a regulatory sequence element with respect to a heterologous gene to be placed under its control is defined by the positioning of the element in its natural setting, i.e. the genes from which it is derived.

Promoter sequences contemplated by the present invention may be native to the host plant to be transformed or may be derived from an alternative source, where the region is functional in the host plant. Other sources include the *Agrobacterium* T-DNA genes, such as the promoters for the biosynthesis of nopaline, octapine, mannopine, or other opine promoters; promoters from plants, such as the ubiquitin promoter; tissue specific promoters (see, e.g. U.S. Patent No. 5,459,252; International Patent Publication No. WO 91/13992); promoters from viruses (including host specific viruses), or partially or wholly synthetic promoters. Numerous promoters that are functional in mono- and dicotyledonous plants are well known in the art (see, for example, Greve, *J. Mol. Appl. Genet. 1:* 499-511, 1983; Salomon *et al., EMBO J. 3:* 141-146, 1984; Garfinkel *et al., Cell 27:* 143-153, 1983; Barker *et al., Plant Mol. Biol. 2:* 235-350, 1983); including various promoters isolated from plants (such as the Ubi promoter from the maize *ubi-1* gene, e.g. U.S. Patent No. 4,962,028) and viruses (such as the cauliflower mosaic virus promoter, CaMV 35S).

In the context of the present invention, a particularly useful tissue-specific promoter is one which drives expression specifically in the stems of sugarcane plants. Such a stem-specific promoter is, for example, that described in International Patent Publication No. WO 01/18211.

The promoter sequences may include regions which regulate transcription, where the regulation involves, for example, chemical or physical repression or induction (e.g. regulation based on metabolites, light, or other physicochemical factors; see, e.g. International Patent Publication No. WO 93/06710 disclosing a nematode responsive promoter) or regulation based on cell differentiation (such as associated with leaves, roots,

seed, or the like in plants; see, e.g. U.S. Patent No. 5,459,252 disclosing a root-specific promoter). Thus, the promoter region, or the regulatory portion of such region, is obtained from an appropriate gene that is so regulated. For example, the ribulose 1,5-bisphosphate carboxylase gene is light-induced and may be used for transcriptional initiation. Other genes are known which are induced by stress, temperature, wounding, pathogen effects, etc.

The chimeric genetic construct of the present invention may also comprise a 3' non-translated sequence. A 3' non-translated sequence refers to that portion of a gene comprising a DNA segment that contains a polyadenylation signal and any other regulatory signals capable of effecting mRNA processing or gene expression. The polyadenylation signal is characterized by effecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. Polyadenylation signals are commonly recognized by the presence of homology to the canonical form 5' AATAAA-3' although variations are not uncommon.

The 3' non-translated regulatory DNA sequence preferably includes from about 50 to 1,000 nucleotide base pairs and may contain plant transcriptional and translational termination sequences in addition to a polyadenylation signal and any other regulatory signals capable of effecting mRNA processing or gene expression. Examples of suitable 3' non-translated sequences are the 3' transcribed non-translated regions containing a polyadenylation signal from the nopaline synthase (*nos*) gene of *Agrobacterium tumefaciens* (Bevan *et al., Nucl. Acid. Res. 11:* 369, 1983) and the terminator for the T7 transcript from the octopine synthase gene of *Agrobacterium tumefaciens*. Alternatively, suitable 3' non-translated sequences may be derived from plant genes such as the 3' end of the protease inhibitor I or II genes from potato or tomato, the soybean storage protein genes and the pea E9 small sub-unit of the ribulose-1,5-bisphosphate carboxylase (ssRUBISCO) gene, although other 3' elements known to those of skill in the art can also be employed. Alternatively, 3' non-translated regulatory sequences can be obtained *de novo* as, for example, described by An (*Methods of Enzymology 153:* 292, 1987), which is incorporated herein by reference.

A genetic construct can also be introduced into a vector, such as a plasmid. Plasmid vectors include additional DNA sequences that provide for easy selection, amplification, and transformation of the expression cassette in prokaryotic and eukaryotic cells, e.g. pUC-derived vectors, pSK-derived vectors, pGEM-derived vectors, pSP-derived vectors, or pBS-derived vectors. Additional DNA sequences include origins of replication to provide for autonomous replication of the vector, selectable marker genes, preferably encoding, for example, antibiotic or herbicide resistance or green fluorescent protein or other visible markers, unique multiple cloning sites providing for multiple sites to insert DNA sequences or genes encoded in the chimeric genetic construct, and sequences that enhance transformation of prokaryotic and eukaryotic cells.

The vector preferably contains an element(s) that permits either stable integration of the vector or a chimeric genetic construct contained therein into the host cell genome, or autonomous replication of the vector in the cell independent of the genome of the cell. The vector, or a construct contained therein, may be integrated into the host cell genome when introduced into a host cell. For integration, the vector may rely on a foreign or endogenous DNA sequence present therein or any other element of the vector for stable integration of the vector into the genome by homologous recombination. Alternatively, the vector may contain additional nucleic acid sequences for directing integration by homologous recombination into the genome of the host cell. The additional nucleic acid sequences enable the vector or a construct contained therein to be integrated into the host cell genome at a precise location in the chromosome. To increase the likelihood of integration at a precise location, the integrational elements should preferably contain a sufficient number of nucleic acids, such as 100 to 1,500 base pairs, preferably 400 to 1,500 base pairs, and most preferably 800 to 1,500 base pairs, which are highly homologous with the corresponding target sequence to enhance the probability of homologous recombination. The integrational elements may be any sequence that is homologous with the target sequence in the genome of the host cell. Furthermore, the integrational elements may be non-encoding or encoding nucleic acid sequences.

For cloning and sub-cloning purposes, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in a host cell such as a bacterial cell. Examples of bacterial origins of replication are the origins of replication of plasmids pBR322, pUC19, pACYC177, and pACYC184 permitting replication in *E. coli*, and pUB110, pE194, pTA1060, and pAMβ1 permitting replication in *Bacillus*. The origin of replication may be one having a mutation to make its function temperature-sensitive in a *Bacillus* cell (see, e.g. Ehrlich, *Proc. Natl. Acad. Sci. USA 75:* 1433, 1978).

To facilitate identification of transformed cells, the vector desirably comprises a further genetic construct comprising a selectable or screenable marker gene. The actual choice of a marker is not crucial as long as it is functional (i.e. selective) in combination with the plant cells of choice. The marker gene and the nucleotide sequence of interest do not have to be linked, since co-transformation of unlinked genes as, for example, described in U.S. Patent No. 4,399,216 is also an efficient process in plant transformation.

Included within the terms selectable or screenable marker genes are genes that encode a "secretable marker" whose secretion can be detected as a means of identifying or selecting for transformed cells. Examples include markers that encode a secretable antigen that can be identified by antibody interaction, or secretable enzymes that can be detected by their catalytic activity. Secretable proteins include, but are not restricted to, proteins that are inserted or trapped in the cell wall (e.g. proteins that include a leader sequence such as that found in the expression unit of extensin or tobacco PR-S); small, diffusible proteins detectable, for example, by ELISA; and small active enzymes detectable in extracellular solution such as, for example, α-amylase, β-lactamase, phosphinothricin acetyltransferase).

Examples of bacterial selectable markers are the *dal* genes from *Bacillus subtilis* or *Bacillus licheniformis*, or markers that confer antibiotic resistance such as ampicillin, kanamycin, erythromycin, chloramphenicol or tetracycline resistance. Exemplary selectable markers for selection of plant transformants include, but are not limited to, a *hyg* gene which encodes hygromycin B resistance; a neomycin phosphotransferase (*npt*) gene conferring resistance to kanamycin, paromomycin, G418 and the like as, for example,

described by Potrykus *et al.* (*Mol. Gene. Genet. 199:* 183, 1985); a glutathione-S-transferase gene from rat liver conferring resistance to glutathione derived herbicides as, for example, described in EP-A 256 223; a glutamine synthetase gene conferring, upon overexpression, resistance to glutamine synthetase inhibitors such as phosphinothricin as, for example, described International Patent Publication No. WO 87/05327, an acetyl transferase gene from *Streptomyces viridochromogenes* conferring resistance to the selective agent phosphinothricin as, for example, described in European Patent Application No. EP-A 275 957, a gene encoding a 5-enolshikimate-3-phosphate synthase (EPSPS) conferring tolerance to N-phosphonomethylglycine as, for example, described by Hinchee *et al.* (*Biotech 6:* 915, 1988), a *bar* gene conferring resistance against bialaphos as, for example, described in International Patent Publication No. WO 91/02071; a nitrilase gene such as *bxn* from *Klebsiella ozaenae* which confers resistance to bromoxynil; a dihydrofolate reductase (DHFR) gene conferring resistance to methotrexate (Thillet *et al.*, *J. Biol. Chem. 263:* 12500, 1988); a mutant acetolactate synthase gene (ALS), which confers resistance to imidazolinone, sulfonylurea or other ALS-inhibiting chemicals (European Patent Application No. EP-A-154 204) or a mutated anthranilate synthase gene that confers resistance to 5-methyl tryptophan.

Preferred screenable markers include, but are not limited to, a *uidA* gene encoding a β-glucuronidase (GUS) enzyme for which various chromogenic substrates are known; a β-galactosidase gene encoding an enzyme for which chromogenic substrates are known; an aequorin gene (Prasher *et al.*, *Biochem. Biophys. Res. Comm. 126:* 1259, 1985), which may be employed in calcium-sensitive bioluminescence detection; a green fluorescent protein gene (Niedz *et al.*, *Plant Cell Reports 14:* 403, 1995); a luciferase (*luc*) gene (Ow *et al.*, *Science 234:* 856, 1986), which allows for bioluminescence detection; a β-lactamase gene (Sutcliffe, *Proc. Natl. Acad. Sci. USA 75:* 3737, 1978), which encodes an enzyme for which various chromogenic substrates are known (e.g. PADAC, a chromogenic cephalosporin); an R-locus gene, encoding a product that regulates the production of anthocyanin pigments (red colour) in plant tissues (Dellaporta *et al.*, in *Chromosome Structure and Function* pp. 263-282, 1988); an α-amylase gene (Ikuta *et al.*, *Biotech 8:* 241, 1990); a tyrosinase gene (Katz *et al*, *J. Gen. Microbiol. 129:* 2703, 1983) which

encodes an enzyme capable of oxidizing tyrosine to dopa and dopaquinone which in turn condenses to form the easily detectable compound melanin; or a *xylE* gene (Zukowsky *et al., Proc. Natl. Acad. Sci. USA 80:* 1101, 1983), which encodes a catechol dioxygenase that can convert chromogenic catechols.

A further aspect of the present invention provides a transfected or transformed cell, tissue, or organ from a C4 grass, which comprises a nucleotide sequence encoding one or more enzymes required for the production of a useful product.

The vectors and chimeric genetic construct(s) of the present invention may be introduced into a cell by various techniques known to those skilled in the art. The technique used may vary depending on the known successful techniques for that particular organism.

Techniques for introducing vectors, chimeric genetic constructs and the like into cells include, but are not limited to, transformation using $CaCl_2$ and variations thereof, direct DNA uptake into protoplasts, PEG-mediated uptake to protoplasts, microparticle bombardment, electroporation, microinjection of DNA, microparticle bombardment of tissue explants or cells, vacuum-infiltration of tissue with nucleic acid, and T-DNA-mediated transfer from *Agrobacterium* to the plant tissue.

For microparticle bombardment of cells, a microparticle is propelled into a cell to produce a transformed cell. Any suitable ballistic cell transformation methodology and apparatus can be used in performing the present invention. Exemplary apparatus and procedures are disclosed by Stomp *et al.* (U.S. Patent No. 5,122,466) and Sanford and Wolf (U.S. Patent No. 4,945,050). When using ballistic transformation procedures, the genetic construct may incorporate a plasmid capable of replicating in the cell to be transformed.

Examples of microparticles suitable for use in such systems include 0.1 to 10 μm and more particularly 0.5 to 5 μm tungsten particles or gold spheres. The DNA construct may be deposited on the microparticle by any suitable technique, such as by precipitation.

Plant tissue capable of subsequent clonal propagation, whether by organogenesis or embryogenesis, may be transformed with a chimeric genetic construct of the present invention and a whole plant generated therefrom. The particular tissue chosen will vary depending on the clonal propagation systems available for, and best suited to, the particular species being transformed. Exemplary tissue targets include leaf disks, embryos, cotyledons, hypocotyls, megagametophytes, callus tissue, existing meristematic tissue (e.g. apical meristem, axillary buds, and root meristems), and induced meristem tissue (e.g. cotyledon meristem and hypocotyl meristem).

The regenerated transformed plants may be propagated by a variety of means, such as by clonal propagation or classical breeding techniques. For example, a first generation (or T1) transformed plant may be selfed to give a homozygous second generation (or T2) transformant, and the T2 plants further propagated through classical breeding techniques.

Even more particularly, the present invention provides a plant cell or multicellular plant or progeny thereof wherein said cell, plant, progeny or part thereof exhibits an activity to manufacture PHAs.

The term "genetically modified" is used in its broadest sense and includes introducing gene(s) into cells, mutating gene(s) in cells and altering or modulating the regulation of gene(s) in cells. In the context of the present invention, a transgenic cell or plant line may also be considered as a mutant cell or plant line when compared with its non-transgenic counterpart. In essence, a selected plant is first genetically modified to introduce a genetic sequence encoding a desired product or intermediate.

Where genetic sequences for more than one gene are to be used in the performance of the present invention, they may be introduced simultaneously or sequentially, separately or together, into the target cells that are to be transformed. For example, a singe genetic construct may comprise all the required genetic sequences for the practice of the subject invention, and this single construct may be introduced into the cells *via* any number of different means, as discussed below. Moreover, each genetic sequence may be operable

linked to and under the control of its own promoter, or may be comprised within a single polycistronic unit. Alternatively, separate genetic constructs may be utilized, each comprising one of the needed genetic sequences. In this event, more than one construct may be introduced into the target cells simultaneously or sequentially. Here and elsewhere throughout the subject specification, the terms "target cells" and "cells to be transformed" should be regarded as being synonymous and refer to cells of a C4 grass that are to be used in accordance with the present invention as a bioreactor.

The one or more genetic sequences, introduced into a C4 grass plant cell, need to be expressed in order to enable the manufacture and accumulation of a product. The term "expression" is to be construed in its broadest sense and includes and encompasses transciption and translation of a genetic sequence to a translation product.

Some plant cells may already comprise a homolog of one or more of the genetic sequences encoding enzymes needed for the production of a given product. In instances where a target plant cell, such as a sugarcane cell, already comprises one or more suitable genetic sequences, capable of directing sufficiently high expression, only those enzymes missing in a given pathway need be provided through *via* a genetic construct as described above.

The present invention extends to homologs and derivatives of any suitable sequences, whether found naturally in a target cell or provided exogenously having been derived from another plant, animal, protist, fungal, archeal or bacterial source, *inter alia*. The derivatives may be at the protein or nucleic acid level.

By "derivative" in relation to a polypeptide is meant a polypeptide that has been derived from the basic sequence by modification, for example, by conjugation or complexing with other chemical moieties or by post-translational modification techniques as would be understood in the art. The term "derivative" also includes within its scope alterations that have been made to a parent sequence including additions, or deletions that provide for functionally-equivalent molecules. Accordingly, the term "derivative" encompasses molecules that affect a plant's phenotype in the same way as does the parent an amino acid

sequence from which it was generated. Also encompassed are polypeptides in which one or more amino acids have been replaced by different amino acids. It is well understood in the art that some amino acids may be changed to others with broadly similar properties without changing the nature of the activity of the polypeptide (conservative substitutions) as described hereinafter. These terms also encompass polypeptides in which one or more amino acids have been added or deleted, or replaced with different amino acids.

"Polypeptide", "peptide" and "an amino acid sequence" are used interchangeably herein to refer to a polymer of amino acid residues and to variants and synthetic analogues thereof. Thus, these terms apply to amino acid polymers in which one or more amino acid residues is a synthetic non-naturally-occurring amino acid, such as a chemical analogue of a corresponding naturally-occurring amino acid, as well as to naturally-occurring amino acid polymers.

The term "derivative" also encompasses fragments. A "fragment", as used herein, means a portion or a part of a full-length parent polypeptide, which retains the activity of the parent polypeptide. As used herein, the term "biologically-active fragment" includes deletion mutants and small peptides, for example, of at least 10, preferably at least 20 and more preferably at least 30 contiguous amino acids, which comprise the above activity. Peptides of this type may be obtained through the application of standard recombinant nucleic acid techniques or synthesized using conventional liquid or solid phase synthesis techniques. For example, reference may be made to solution synthesis or solid phase synthesis as described, for example, in Chapter 9 entitled *"Peptide Synthesis"* by Atherton and Shephard which is included in a publication entitled *"Synthetic Vaccines"* edited by Nicholson and published by Blackwell Scientific Publications. Alternatively, peptides can be produced by digestion of an amino acid sequence of the invention with proteinases such as endoLys-C, endoArg-C, endoGlu-C and staphylococcus V8-protease. The digested fragments can be purified by, for example, high performance liquid chromatographic (HPLC) techniques. Any such fragment, irrespective of its means of generation, is to be understood to be encompassed by the term "derivative" as used herein.

The terms "variant" and "homolog" refer to nucleotide sequences displaying substantial sequence identity with a reference nucleotide sequences or polynucleotides that hybridize with a reference sequence under stringency conditions that are defined hereinafter. The terms "nucleotide sequence", "polynucleotide" and "nucleic acid molecule" may be used herein interchangeably and encompass polynucleotides in which one or more nucleotides have been added or deleted, or replaced with different nucleotides. In this regard, it is well understood in the art that certain alterations inclusive of mutations, additions, deletions and substitutions can be made to a reference nucleotide sequence whereby the altered polynucleotide retains the biological function or activity of the reference polynucleotide. The term "variant" also includes naturally-occurring allelic variants.

The extent of homology may be determined using sequence comparison programs such as GAP. In this way, sequences of a similar or substantially different length to those cited herein might be compared by insertion of gaps into the alignment, such gaps being determined, for example, by the comparison algorithm used by GAP, as is further discussed below.

Homologous sequences will generally hybridize under particular specified conditions. The term "hybridization" denotes the pairing of complementary nucleotide sequences to produce a DNA-DNA hybrid or a DNA-RNA hybrid. Complementary base sequences are those sequences that are related by the base-pairing rules. In DNA-DNA hybridization, A pairs with T and C pairs with G. In DNA-RNA hybridization, U pairs with A and C pairs with G. In this regard, the terms "match" and "mismatch" as used herein refer to the hybridization potential of paired nucleotides in complementary nucleic acid strands. Matched nucleotides hybridize efficiently, such as the classical A-T and G-C base pair mentioned above. Mismatches are other combinations of nucleotides that do not hybridize efficiently.

The extent of hybridization that may be displayed by homologous sequences depends on the conditions of, for example, temperature, ionic strength, presence or absence of certain organic solvents, under which hybridization and washing procedures are carried out. The

higher the stringency, the higher will be the degree of complementarity between immobilised target nucleotide sequences and the labelled probe polynucleotide sequences that remain hybridized to the target after washing. "High stringency conditions" refers to temperature and ionic conditions under which only nucleotide sequences having a high frequency of complementary bases will hybridize. The stringency required is nucleotide-sequence dependent, and further depends upon the various components present during hybridization and subsequent washes, and the time allowed for these processes. Generally, in order to maximize the hybridization rate, relatively low-stringency hybridization conditions are selected: about 20 to 25°C lower than the thermal melting point (T_m). The T_m is the temperature at which 50% of specific target sequence hybridizes to a perfectly complementary probe in solution at a defined ionic strength and pH. Generally, in order to require at least about 85% nucleotide complementarity of hybridized sequences, highly stringent washing conditions are selected to be about 5 to 15°C lower than the T_m. In order to require at least about 70% nucleotide complementarity of hybridized sequences, moderately-stringent washing conditions are selected to be about 15 to 30°C lower than the T_m. Highly permissive (very low stringency) washing conditions may be as low as 50°C below the T_m, allowing a high level of mis-matching between hybridized sequences. Those skilled in the art will recognize that other physical and chemical parameters in the hybridization and wash stages can also be altered to affect the outcome of a detectable hybridization signal from a specific level of homology between target and probe sequences.

Reference herein to "low stringency conditions" is generally determined at 42°C and includes and encompasses from at least about 0% v/v to at least about 15% v/v formamide, and from at least about 1 M to at least about 2 M salt for hybridization, and at least about 1 M to at least about 2 M for washing conditions. Alternative stringency conditions may be applied where necessary, such as: medium stringency, which includes and encompasses from at least about 16% v/v to at least about 30% v/v formamide, and from at least about 0.5 M to at least about 0.9 M salt for hybridization, and at least about 0.5 M to at least about 0.9 M salt for washing conditions, or high stringency, which includes and encompasses from at least about 31% v/v to at least about 50% v/v formamide, and from at

least about 0.01 M to least about 0.15 M salt for hybridization, and at least about 0.01 M to at least about 0.15 M salt for washing conditions.

Terms used to describe sequence relationships between two or more nucleotide sequences or amino acid sequences include "reference sequence", "comparison window", "sequence identity", "percentage of sequence identity" and "substantial identity". A "reference sequence" is at least 12 but frequently 15 to 18 and often at least 25 monomer units, inclusive of nucleotides and amino acid residues, in length. Because two polynucleotides may each comprise (1) a sequence (i.e. only a portion of the complete polynucleotide sequence) that is similar between the two polynucleotides, and (2) a sequence that is divergent between the two polynucleotides, sequence comparisons between two (or more) polynucleotides are typically performed by comparing sequences of the two polynucleotides over a "comparison window" to identify and compare local regions of sequence similarity. A "comparison window" refers to a conceptual segment of at least 6 contiguous positions, usually about 50 to about 100, more usually about 100 to about 150 in which a sequence is compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. The comparison window may comprise additions or deletions (i.e. gaps) of about 20% or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Optimal alignment of sequences for aligning a comparison window may be conducted by computerized implementations of algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Drive Madison, WI, USA) or by inspection and the best alignment (i.e. resulting in the highest percentage homology over the comparison window) generated by any of the various methods selected. Reference also may be made to the BLAST family of programs as for example disclosed by Altschul *et al.*, *Nucl. Acids Res. 25:* 3389, 1997. A detailed discussion of sequence analysis can be found in Unit 19.3 of Ausubel *et al.*, "Current Protocols in Molecular Biology" John Wiley & Sons Inc, 1994-1998, Chapter 15.

The term "sequence identity" as used herein refers to the extent that sequences are

identical on a nucleotide-by-nucleotide basis or an amino acid-by-amino acid basis over a window of comparison. Thus, a "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (e.g. A, T, C, G, I) or the identical amino acid residue (e.g. Ala, Pro, Ser, Thr, Gly, Val, Leu, Ile, Phe, Tyr, Trp, Lys, Arg, His, Asp, Glu, Asn, Gln, Cys and Met) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (i.e. the window size), and multiplying the result by 100 to yield the percentage of sequence identity. For the purposes of the present invention, "sequence identity" will be understood to mean the "match percentage" calculated by the DNASIS computer program (Version 2.5 for windows; available from Hitachi Software engineering Co., Ltd., South San Francisco, California, USA) using standard defaults as used in the reference manual accompanying the software.

The one or more constructs may be introduced into a plant cell by any number of well-recognized means such as discussed above.

Preferably, the genetic constructs of the present invention are introduced *via* the use of biolistics.

Accordingly, another aspect of the present invention provides a transgenic C4 grass, cells of which have been transformed with one or more genetic sequences such that one of the following products in the cytosol, storage vacuole, non-plastid organelle or extra-cellular matrix of said cells:

(i) polyhydroxyalkanoates

(ii) vanillin

(iii) sorbitol

(iv) indigo

(v) fructans

(vi) lactic acid

(vii) adipic acid

(viii) 1,3-propanediol

(ix) 2-phenylethanol

(x) pHBA

The present invention extends to parts of plants tissue including leaves, stems, vascular bundles, bark, reproductive material, roots and any extracted liquid ("juice") from said plant.

While the present invention is exemplified using the compounds hereinbefore described, it is to be understood that the invention extends to and encompasses the use of any suitable genetic sequence capable of effecting the production of any product in the cells or extracellular matrix of a C4 grass.

The term "gene" is used in its broadest sense and includes cDNA corresponding to the exons of a gene. Accordingly, reference herein to a "gene" is to be taken to include:-

(i) a classical genomic gene consisting of transcriptional and/or translational regulatory sequences and/or a coding region and/or non-translated sequences (i.e. introns, 5'- and 3'- untranslated sequences); or

(ii) mRNA or cDNA corresponding to the coding regions (i.e. exons) and 5'- and 3'-untranslated sequences of the gene.

The term "gene" is also used to describe synthetic or fusion molecules encoding all or part of an expression product. In particular embodiments, the term "nucleic acid molecule" and "gene" may be used interchangeably.

In order to improve the efficiency and accumulation rate of a product, a suitable genetic sequence or sequences may be more specifically targeted so as to facilitate generation of expression products in particular sub-cellular areas or organelles within the plant. These include, for example, the cytosol, a storage vacuole or a plastid or non-plastid organelle.

The usefulness of a given sub-cellular compartment for a given product depends on the nature and potential toxicity of the product to the plant. For example, PHA production is dependent on both the types of polymer produced and the metabolic pathways being engineered.

One particularly useful sub-cellular area for the production of products such as PHB, 1,3-propanediol and sorbitol, is the cytosol, wherein sucrose is both synthesized *via* gluconeogenesis and degraded *via* glycolysis, leading to the production of pyruvate. In the cytosol, PHB is the preferred polymer, as moderate amounts of acetyl-CoA are available for *pha*A, *pha*B and *pha*C. A particularly useful sub-cellular organelle is the mitochondrion, wherein acetyl-CoA, which may be produced from pyruvate *via* mitochondrial pyruvate dehydrogenase, and/or perhaps from fatty acids *via* β-oxidation, is used to fuel the TCA cycle. A second useful sub-cellular organelle, with moderate to high acetyl-CoA, is the peroxisome, the site of fatty acid degradation *via* β-oxidation. These pathways involve the utilization of substantial amounts of acetyl-CoA, depleting reserves and rendering it unavailable for use in manufacture of a product. However, the pyruvate needed for acetyl-CoA production is generated *via* glycolysis, which, in a sink tissue such as sugarcane stems, is fuelled by sucrose. Hence, the carbon drain that usually results from effecting the production of a product, such as a PHA, vanillin and the like, in a plant cell,

is able to be overcome by the sucrose-accumulating plant cell's ability to mobilise its substantial sucrose stores. The deleterious effects resulting from product accumulation observed in non-sucrose-accumulating plant species does not occur in sugarcane, as a concomitant state of general starvation is precluded by the mobilization of sucrose from storage vacuoles, which replenishes the reduction of cellular acetyl-CoA pools caused by the introduced genetic sequences.

Accumulation of PHAs, and other products (particularly products such as pHBA, adipic acid and indigo) may also be targeted to a plastid, such as a chloroplast, where large amounts of acetyl-CoA are used for fatty acid biosynthesis. For example, for the production of PHAs other than PHB, plastids and peroxisomes are the preferred sub-cellular compartment, as PhaG (plastid) and PhaJ (peroxisome) provide monomers suitable for MCL-PHA polymerases such as PhaC1 from intermediates in fatty acid biosynthesis and β-oxidation. In addition, particular biosynthetic pathways from which a particular product may be derived may exist only in the plastid. For example, the shikimate pathway is locatized in the plastid in plants. Products such as indigo and adipic acid may be derived from intermediates of the shikimate pathway *via* the addition of new biosynthetic enzymes. For these new enzymes to produce the product of interest, they must be localised to the particular organelles where their substrates would be found.

In order to direct product accumulation to a desired sub-cellular location, particular specific "target sequences" may be incorporated into the genetic constructs described above.

A target sequence includes a signal sequence such as a signal sequence to direct the protein to a plastid, vacuole, mitochondrion or other appropriate organ or tissue.

Preferably, accumulation of PHA is in the cytosol or mitochondrion, assisted *via* mobilization of sucrose reserves located in the storage vacuoles of sugarcane stem cells.

Preferably, accumulation of adipic acid and indigo is in the plastid, wherein the introduced

biosynthetic enzymes have access to intermediates of the shikimate pathway.

The plants of the present invention may also be further "tagged" with a reporter that identifies the plant as a plant bioreactor. Any number of physiological or genetic tags would be suitable, and readily identified by one of skill in the art. Examples of physiological "tags" that could be introduced include marker genes such as the green fluorescent protein gene, the firefly luciferase gene and the GUS gene.

Marker genes that alter the physical appearance of the plant may also be used as identifying tags. Examples include increased or decreased length of stems and/or alterations to color. Furthermore, a number of resistance phenotypes may also be used to identify the plant bioreactors. Genes encoding resistance to pests such as bacterial, fungal or nematode pests have been identified in the art, and would be suitable as "tags".

In addtion, a genetic sequence itself may comprise the tag, referred to herein as "DNA barcoding". Tagging in this manner is done by introducing a known non-coding polynucleotide sequence into the plant. The tag may then be amplified from the plant using known PCR primers. Plants may then be identified as bioreactors according to the present invention by the presence of a particular size amplicon after the PCR reaction. Further discrimination, for example between types of plant bioreactor, may be achieved by altering the polynucleotide sequence of the DNA barcode in the region between the PCR primers. In this way, the sequence of the bacrocode may be elucidated using automated sequencing techniques to determine the exact identity of the plant. This technique allows for a generic test to identify all plant bioreactors, and allows further discrimination to identify the type of bioreactor based on the sequence of the DNA barcode.

The genetic sequences comprising or encoding the "tag" may be introduced to the plant using the methods hereinbefore described. The tag may be introduced on the same construct as the biosynthetic gene, or may be independently introduced. If introduced independently, the tage may be introduced on a different construct at the same time as transformation with the biosynthetic gene, or introduced to the plant before or after the

biosynthetic gene.

Accordingly, the present invention contemplates a plant suitable for use as a bioreactor that has been tagged with a genetic sequence which encodes or comprises a genotypic or phenotypic feature that allows differentiation of the plant bioreactor from a wild-type plant.

Preferably the plant is a C4 grass, and more preferrably, the plant is sugarcane.

The plant-based bioreactor system of the present invention is useful in enabling the production of molecules such as PHAs, pHBA, vanillin, sorbitol, indigo, fructans, lactic acid, adipic acid, 1,3-propanediol, 2-phenylethanol, *inter alia*, by a number of different parties such as different commercial entities. The present invention extends, therefore, to a data processing system to monitor the use of the plants and/or the production of target molecules.

Accordingly, another aspect of the present invention contemplates a method for generating a target molecule in a sucrose-accumulating plant, said method comprising:-

(i) providing a plant or cells of a plant to a party; and

(ii) permitting the party to generate and harvest molecules from said plant or cells of said plant receiving and processing data from said party.

The data received from the party includes, for example, numbers of plants grown and/or harvested, the types of genetic constructs introduced into the cells and/or income received from sale of the products.

The present invention is further described by the following non-limiting Examples.

EXAMPLE 1

Materials

Restriction digests, DNA ligations and all other DNA manipulations were performed as described in Sambrook, *et al.,* Molecular Cloning, A Laboratory Manual, 2nd edition, Cold Spring Harbor Press, 1989.

EXAMPLE 2

*Cloning of the pha*C1 *gene from P. aeruginosa*

The phaC1 gene targeted to plant peroxisomes inserted into pART27 as an *Eco*RI/*Xba*I fragment, was obtained from Y. Poirier (University of Lausanne, Switzerland). In order to express the gene in sugarcane, it was excised with the said enzymes, end-filled with T4 DNA polymerase (Promega) and inserted into the *Sma*I site of pUBI-MCS-Nos. To achieve targeting of the *pha*C1 gene product to mitochondria and plastids, the gene is modified as described below.

EXAMPLE 3

Generation of genetic constructs

(a) Constructs comprising sequences encoding PHB-synthesizing enzymes

Constructs containing the *pha*A, *pha*B and *pha*C genes from *Ralstonia eutropha* targeted to plastids and cloned in pUC 18 as *Xba*I-*Sac*I fragments were obtained from Y. Poirier (University of Lausanne, Switzerland).

The *pha*A, *pha*B and *pha*C genes derived from *Ralstonia eutropha* were cloned into the vector pU3z. This vector is a derivative of pGEM3 (Promega) containing the maize polyubiquitin promoter and *nos* terminator from *A. tumefaciens*, and works well as an

expression vector in sugarcane. All genes were amplified/modified using the polymerase chain reaction (PCR) prior to insertion into pU3z except *pha*C1, which was blunt-end cloned into the same vector. All constructs used for plant transformation are listed below. Where inserts are modified, they are sequenced in full to ensure quality.

PCR modifications were performed as follows. *Platinum Pfx* (registered trademark) DNA polymerase, 10x buffer and PCR-enhancer were obtained from Invitrogen. Final concentrations were one unit of polymerase per reaction, 1x buffer and enhancer, 2 mM Mg^{2+}, 0.2 mM dNTP, 0.4 μM of each primer. All primers were purchased from Geneworks, Australia. Reactions were performed in a MJC PTC-100 thermal cycler. The profile was initial denaturation at 96°C for 5 min, followed by 35 cycles of 94°C for 30 seconds, 42°C for 30 sec and 72°C for 2.5 min. a final extension step of 72°C for 10 min preceded a final hold at 4°C. Table 1 lists primers used in PCR reactions.

TABLE 1

NAME	SEQUENCE	POSITION, GENE, TARGET
TphaF	N_6ggatccatggcttctatgatatcct [SEQ ID NO:34]	5', phaA-C, plastid
PhaF	N_6GGATCCATGACTGACGTTGTCATC [SEQ ID NO:35]	5', phaA, cytosol
PhbF	N_6GGATCCATGACTCAGCGCATTGCG [SEQ ID NO:36]	5', phaB, cytosol
PhcF	N_6GGATCCATGGCGACCGGCAAAGGC [SEQ ID NO:37]	5', phaC, cytosol
PhaR	CTGAGTCATGTCCACTCC [SEQ ID NO:38]	3', phaA, cytosol and plastid
PhbR	CTGCCGACTGGTGGAACC [SEQ ID NO:39]	3', phaB, cytosol and plastid
PhcR	GAAGCGTCATGCCTTGGC [SEQ ID NO:40]	3', phaC, cytosol and plastid
PhaC1CF	N_6GGATCCATGAGCCAGAAGAAC [SEQ ID NO:41]	5', phaC1, cytosol and mitochondia
PhaC1CR	N_6GGTACCTCATCGTTCATGCACG [SEQ ID NO:42]	3', phaC1, cytosol and

		plastid
PhaC1PF	N_6CCCGGGTGAGCCAGAAGAACAATAAC [SEQ ID NO:43]	5', phaC1, plastid
PhaJF	GGATCCATGAGCGCACAATCCCTGG [SEQ ID NO:44]	5', phaJ, peroxisome
PhaJR	AAGCTTTTGAAGGCAGCTTGACCACGGC [SEQ ID NO:45]	3', phaJ, peroxisome
PhaGF	CCCGGGTGAGGCCAGAAATCGCTGTAC [SEQ ID NO:46]	5', phaG, plastid
PhaGR	GGTACCTCAGATGGCAAATGCATGC [SEQ ID NO:47]	3', phaG, plastid
SSP-F	NNGAGCTCGATGGGAGGTGCTCGAAGACATATTACC [SEQ ID NO:48]	5', stem-specific promoter
SSP-R	NNGGATCCTGTACTAGATATGGCAGC [SEQ ID NO:49]	3', stem-specific promoter

Approximately 10 ng of construct was used as template in each reaction.

Following PCR, fragments were gel purified and cloned into the *BamHI* and *Sma*I sites of pUSN in the correct orientation between the maize polyubiquitin promoter and the *nos*-terminator from *A. tumefaciens*. Plasmid constructs were fully sequenced and purified by anion-exchange chromatography (Qiagen, Australia) prior to transformation into callus tissue.

(b) Constructs comprising marker sequences

Two plasmids were obtained from the CSIRO, Brisbane, Australia. The first comprised the <pUbi-gfp-nos> construct, which carries the green fluorescent protein (GFP) from *Aquorea victoria* under the control of the same promoter as above. The second plasmid, designated "pEmuKn", harbours an *aphA* gene (neomycin phosphotransferase) under the control of the Emu promoter. These plasmids were used without any further modification.

EXAMPLE 4

Sugarcane transformation

(a) Generation of embryogenic callus tissue

Embryogenic callus of the sugarcane variety Q117 was established, as described in Bower *et al., Molec. Breeding 2:* 239-249, 1996. Briefly, embryogenic callus was established by excision of inner leaf whorls from cane tops 2-5 cm above the apical meristem. Disks of approximately 2 mm thickness were placed on MSC_3 medium that contains 3 µg/ml 2,4-Dichlorophenoxy-acetic acid (2,4-D) and incubated at 28°C in the dark for 2-5 months, with fortnightly subculturing. In order to avoid problems associated with stress arising from the tissue culture process, such as somaclonal variation, unused callus tissue was discarded after 6 months in culture.

Prior to transformation, embryogenic callus was transferred to osmotic MSC_3, as previously described (Bower *et al.*, 1996, *supra*)

(b) Bombardment of callus tissue

DNA was coated onto tungsten particles (Sylvania M-10) and embryogenic callus bombarded as described by Bower *et al.* (1996, *supra*).

Callus tissue was co-transformed with up to five individual constructs, plasmid solutions being mixed to give equimolar concentrations to facilitate co-integration and expression of genes required for PHB production. There is a strong correlation between co-transformation and co-integration of constructs into the genome of plant hosts. The combinations of constructs introduced into Q117 callus and the target for products of each *pha* gene are shown in the following Table 2, wherein "cyt" = cytosol, "pla" = plastid, "mito" = mitochondrion and "perox" = peroxisome:

TABLE 2

COMBINATION AND TARGETING OF *pha* GENE PRODUCTS
phaA (cyt), phaB (cyt), phaC (cyt), Ubi-GFP, Emu-Kn
phaA (pla), phaB (pla), phaC (pla), Ubi-GFP, Emu-Kn
phaA (cyt), phaB (cyt), phaC1 (cyt), Ubi-GFP, Emu-Kn
phaC1 (perox), Ub-GFP, Emu-Kn
phaA (cyt), phaB (cyt), phaC (cyt), phaC1 (perox), phaA (pla), phaB (pla), phaC (pla), phaG (pla), phaC1 (pla), Ubi-GFP, Emu-Kn
phaC1 (perox), phaA (pla), phaB (pla), phaC (pla), phaG (pla), phaC1 (pla), Ubi-GFP, Emu-Kn
phaA (pla), phaB (pla), phaC (pla), phaG (pla), phaC1 (pla), Ubi-GFP, Emu-Kn
phaG (pla), phaC1 (pla), Ubi-GFP, Emu-Kn
phaA (mito), phaB (mito), phaC (mito), Ubi-GFP, Emu-Kn

Microprojectile bombardment was performed as described in Bower *et al.* (*supra*), except that, to improve transformation efficiencies, the vacuum chamber was evacuated to -100 kPa atmospheric pressure and particles accelerated by a helium pulse of 3000 kPa for 100 ms.

(c) Selection of transformed material

Following bombardment, callus was allowed to recover for one hour, before being placed onto MSC_3 medium supplemented with 50 μg/ml Geneticin (registered trademark) (Invitrogen).

Putatively transformed tissue was screened for the expression of GFP and resistance to the antibiotic geneticin. Bombarded callus was examined for the presence of cells expressing GFP 7 days after transformation and stained *in vivo* with Nile Red, a sensitive *in vivo* stain specific for intracellular lipids, as previously described (Taguchi *et al., FEMS Microbiol.*

Lett. 198: 65-71, 2001; Greenspan, *et al.,* 1985, *supra*). For both techniques, an Olympus SZX 12 stereomicroscope equipped with GFP excitation and emission filters was used. Dump

Antibiotic selection was continued for 3 months in the dark at 28°C and continued during plant regeneration. Only calli expressing both selectable markers were allowed to regenerate into plantlets.

(d) Regeneration of transformed sugarcane plantlets

For plant regeneration, callus was transferred to medium without 2,4-D and incubated at 24-26°C under illumination. Plantlets appeared after 2-4 months and were transferred to potting mix and kept in mini-glasshouse (Yates, Australia) for one week prior to transfer to glasshouse facilities.

EXAMPLE 5

Determination of quantity and composition of PHA produced

If the desired PHA is PHB, quantification in transgenic sugarcane is conducted by HPLC, as described by Karr *et al., Applied and Environmental Microbiology 46:* 1339-1344, 1983. This method allows for the analysis of plant extracts with minimal handling. This technique is illustrated in Figure *.

For other PHAs, molecular characterization and quantification of PHA content in transgenic plants is carried out using gas chromatography analysis.

For GC analysis, PHA was separated from homogenized leaf samples by chloroform extraction, followed by methanol extraction, to remove lipids other than PHA. The polymer was then purified further by acetone extraction.

Transesterification of plant extracts was performed as described in Braunegg *et al., Eur. J. Appl. Microbiol. Biotechnol. 6:* 29-37, 1978 modified by using boron trifluoride rather than sulfuric acid as catalyst and decreased incubation time to 1 hr. Gas chromatography analysis was performed using a Varian 3300 chromatograph, as described by Slater *et al.,* in *J. Bacteriol. 180:* 667-73, 1998. Purified PHB (Coparsucar, Brazil) and methyl-3-hydroxybutyrate (Sigma) were used as positive controls either pure or spiked into negative control plant extracts. Modifications to this method, namely using Boron trifluoride rather than sulfuric acid as catalyst and decreasing incubation times to 1 hour, in addition to minor modifications to instrument parameters, were found to improve peak resolution.

Millenium software (Waters Corp., Milford, MA) is used to quantify amounts of PHAs produced, by comparison of peak areas from plant extracts with those from standards of known concentration. GC-MS is used to determine the composition of PHAs produced in transgenic sugarcane, as different hydroxy-alkanoates have different mass-spectrum signatures.

EXAMPLE 6

Production of PHB in transgenic sugarcane leaves

In plants accumulating high amounts of PHA, gene copy number, transcription levels and amount of protein are determined using standard molecular biology techniques. Gene copy numbers are determined by Southern blot analysis of genomic DNA from transformed plants producing PHAs. Transcription levels will be evaluated by northern blot analysis of RNA from transgenic plants and gene product levels examined by western blot analysis of protein extracts using antibodies against the gene products. The antisera were obtained from Prof. Y. Poirier (University of Lausanne, Switzerland).

EXAMPLE 7

Targeting of PHB production to non-plastid sugarcane organelles

The constructs pUbi-phaA, pUbi-phaB and pUbi-phaC comprising the *pha*A, B and C genes, respectively, transcriptionally fused to the aforementioned maize polyubiquitin promoter were digested with *Bam*HI and dephosphorylated with shrimp alkaline phosphatase (Promega, Maddison, USA). A *Bgl*II/*Bam*HI fragment of the plasmid sB-pma4-35S-β-del-GFP, containing the leader sequence and first 12 amino acid residues of the β subunit of the *Nicotiana plumbaginifolia* mitochondrial F1-ATPase (Chaumont *et al., PMB 24:* 631-664, 1994), was ligated with pUbi-phaA, B or C cut with *Bam*HI. These constructs target PhaA, B or C to the mitochondria with high efficiency. For targeting of pha gene products to other organelles, the genes were modified and ligated into plasmids already containing the required DNA targeting sequences.

Sugarcane was transformed with these constructs, using the aforementioned methods.

Transgenic plants thereby generated are screened for PHB production, using the aforementioned techniques.

EXAMPLE 8

Production of PHB in transgenic sugarcane stems

Plasmids pUbi-phaA, pUbi-phaB, pUbi-phaC, pUi-TP-phaA, pUbi-TP-phaB and pUbi-TP-phaC were digested with *Bam*HI/*Eco*RI to release the *pha* genes with or without the aforementioned plastid leader sequence at the 5' end, and with the aforementioned NOS terminator at the 3' end. These fragments were ligated into the vector pSSP cut with *Bam*HI/*Eco*RI. pSSP is a derivative of the vector p67G-420 (supplied by Prof. R Birch, University of Queensland). p67G-420 houses, immediately upstream of a unique *Bam*HI site, a stem-specific promoter isolated from sugarcane. To obtain pSSP, a consensus ribosome binding site was removed from the 3' end of the promoter in the following way.

The promoter was PCR amplified, using the aformentioned technique, with the primers SPP-F and SPP-R (see Table 1), incorporating *Bam*HI and *Sac*I sites into the 5' and 3' ends of the promoter, repectively. The PCR product was digested with *Bam*HI and *Sac*I and then religated into the backbone of p67G_420 digested with the same enzymes. The promoter was fully sequenced to confirm quality. The resulting constructs, which were fully sequenced to confirm quality, drive gene expression in the stems and target the gene products to either the cytosol or the plastids. pSSP was linearized with *Bam*HI and dephosphorylated with shrimp alkaline phosphatase, and ligated with the aforementioned *Bgl*II/*Bam*HI fragment of the plasmid sB-pma4-35S-β-del-GFP, giving the intermediate vector pSSP-Tm. *Bam*HI/*Eco*RI fragments of pUbi-phaA, pUbi-phaB and pUbi-phaC were ligated into pSSP-Tm. The resulting constructs, which were fully sequenced to confirm quality, drive gene expression in the stems and target the gene products to the mitochondria.

Sugarcane was tranformed with these constructs using the aforementioned methods.

Transgenic plants thereby generated are screened for PHB production, using the aforementioned techniques.

EXAMPLE 9

Detection of PHB in chloroplasts of transgenic sugarcane

Transgenic plants expressing PHB biosynthetic genes were produced according to the methods described herein. Accumulation of PHB in the plastid was assessed using both HPLC and transmission electron microscopy.

Figure * shows a graphical representation of the detection of PHB in chloroplasts of transgenic sugarcane. Panels A-C indicate detection of PHB using HPLC. Panel A is a wild-type sugarcane control; Panel B is the plant in A spiked with PHB; Panel C depicts a transgenic sugarcane line accumulating PHB in the plastids. Arrows point to the elution

point of crotonic acid, which is the product of acid-hydrolsis of PHB. The insert in Panel C shows that the peak at 30 min in C has the same spectrum as crotonic acid.

Panels D-F in Figure 4 show the detection of PHB granules in plants by transmission electron microscopy. Panel D shows a positive control comprising a chloroplast from a mesophyll cell in a PHB +ve Arabidopsis plant (Bohmert *et al.* 2000). Panel E is a electron-micrograph showing PHB granules in a chloroplast of a mesophyll cell from a PHB-producing sugarcane plant. Panel F shows PHB granules in a chloroplast of a bundle-sheath cell from the same plant line in E. Scale bars = 200 nm.

EXAMPLE 10

Agronomic performance of PHB-producing sugarcane lines

Four transgenic sugarcane lines expressing the PHB biosynthesis genes of *Ralstonia eutropha* were grown for 3 months in a randomised glasshouse plot. Control plants comprised GFP-expressing and tissue-culture-regenerated wild-type plants. PHB content was assessed in lamina from the tips of mature leaves and quantified by HPLC analysis.

The results are shown in Figure 5. The production of PHB in sugarcane at up to 1.6% of leaf dry-weight did not reduce agronomic performance compared with GFP-expressing and wild-type control plants. Data are the mean ± SE ($n = 3$). DW = dry-weight.

EXAMPLE 11

Affect of PHB production on sugarcane sugar accumulation.

The plants assessed in Example 10, ie. PHB producing, GFP expressing and wild-type sugarcane, were further examined for their sucrose, glucose and fructose concentrations to determine the effect of PHB production on sugar content.

The results are shown in Figure 6. It was observed that PHB accumulation of up to 1.6% of leaf dry-weight did not reduce sucrose, glucose, fructose or total sugar content in PHB

producing (solid bars) plants compared to GFP-expressing (open bars) and wild-type (hatched bars) controls. Data are the mean ± SE ($n = 3$). DW = dry-weight.

EXAMPLE 12

Distribution of PHB in PHB-producing sugarcane

The distribution of PHB throughout transgenic sugarcane line PHB3 was determined by HPLC analysis. Samples were taken from:

(i) lamina of the tip, midpoint and base of young, intermediate and mature leaves;

(ii) combined rind and pith of young, intermediate and mature stem internodes; and

(iii) roots.

The PHB content data are presented as the mean percentage of leaf dry-weight ± SE (n = 3). ND = not detected.

EXAMPLE 13

Production of Vanillin in Sugarcane

Vanillin (4-hydroxy-3-methoxybenzaldehyde) would be produced as a co-product with sucrose. Sucrose yield is expected to decrease in direct proportion to the amount of vanillin produced.

Initially, genes for the vanillin biosynthetic pathway from a known source are cloned. These genes are then expressed in sugarcane, including any tailoring of the expression pattern as required. The product is produced as a glucose conjugate, which is stable.

A number of biological pathways have been discovered for the biosynthesis/biodegradation of vanillin. At least 2 of these have substrates which are available in plants.

1) 3-Dehydroshikimic acid is produced as an intermediate in the shikimate pathway. A pathway has been identified which converts this substrate *via* 3-Dehydroshikimate dehydratase to protocatechuic acid then to vanillic acid *via* Catechol-o-methyltransferase and finally to vanillin *via* Aryl aldehyde dehydrogenase.

2) Ferulic acid is a secondary metabolite of the phenylpropanoid pathway involved in lignin synthesis. It is converted *in planta* to feruloyl-CoA by feruloyl-CoA synthetase which in turn is converted to vanillin by enoyl-CoA hydratase/aldolase.

Glucosylation of the product in vivo is expected to detoxify the product. Accordingly, inducible expression should not be required. The maxiumum level of production is determined by the flux through the phenylpropanoid pathway. However, Sugarcane has a productive phenylpropanoid pathway and should adapt readily to increased demands placed on it for synthesis of vanillin.

EXAMPLE 14

Production of Sorbitol in Sugarcane

Zymomonas mobilis is able to produce sorbitol from sucrose or a mixture of glucose and fructose in a one-step reaction catalysed by the glucose-fructose oxidoreductase GFOR (Genbank accession no. Z80356, M97379). The glucose is oxidized to gluconolactone while the fructose is reduced to sorbitol.

glucose + fructose → sorbitol + gluconolactone

Sorbitol production in sugarcane could be achieved by using GFOR. This involves constructing an expression cassette by fusing GFOR to the maize polyubiquitin promoter and nopaline synthase terminator and introducing the cassette into sugarcane callus by biolistic transformaton. The *Z. mobilis* GFOR is not membrane-bound and resides in the periplasm and should work equally well as a cytosolic enzyme in sugarcane.

Sorbitol production is unlikely to be toxic in sugarcane since sorbitol is found in numerous fruits (apples, pears, plums, berries, cherries). Sorbitol functions physiologically to regulate osmotic stress hence extremely high levels may be detrimental. Vacuolar storage may circumvent this problem.

The threshold level at which sorbitol is deleterious to the host may be determined by growing sugarcane callus on solid medium containing sorbitol.

A potential large-scale system for the recovery of sorbitol from sugarcane involves adding an aqueous organic salt solution, mixing and then separating a salt water phase from a polyol-rich phase (see international patent application WO210252).

EXAMPLE 15

Indigo production in Sugarcane

The chief incentive to use sugarcane as an indigo biofactory is to provide a manufacturing route that will produce relatively inexpensive indigo from a renewable feedstock.

Indigo production by microbial fermentation has been demonstrated by expressing the genes that mediate indigo formation in *E. coli* (Drewlo 2001, Berry 2002). The pigment is derived by converting endogenous tryptophan to indole using the *Enterobacter aerogenes* tryptophanase or L-tryptophan indole lyase EC 4.1.99.1 (Genbank accession no. D14297). Subsequently the indole is converted to indigo *via* two possible reactions.

Route A: Pseudomonas putida napthalene dioxygenase (Genbank accession no. M83949)

Route B: Ralstonia eutropha bec gene (Genbank accession no. AF306552)

Indigo production in sugarcane would involve constructing an expression cassette by fusing the aforementioned genes to the maize polyubiquitin promoter and nopaline synthase terminator and introducing the cassette into sugarcane callus by biolistic transformaton. Both route A and B should be tested if possible. Tryptophan is a product of the plant shikimate pathway, which is responsible for synthesizing lignin precursors. The cloned genes will need to be plastid-targeted since the shikimate pathway reactions reside in this compartment. The available metabolic flux in this pathway is expected to be significant.

EXAMPLE 16

Production of Fructans in Sugarcane

Naturally occurring fructans may contain 10 to 100,000 fructose residues. Bacteria produce the larger fructans whilst those occurring in plants are smaller. The larger polymers are desirable because they are less soluble in water and consequently easier to extract. Larger fructans will not affect the osmotic pressure in the cell to the same degree as smaller molecules. Therefore it is possible to store greater quantities of fructan before the cell is affected.

Numerous bacterial fructosyltransferases or levansucrases have been characterized *(Genbank accession no. AY150365, Bacillus subtilis)*. These enzymes catalyze the transfer of the D-fructosyl residue from sucrose to the β-2,6-linked residues of fructan.

Sucrose → fructan + glucose

Fructan production in sugarcane would be achieved by constructing an expression cassette containing levansucrase, the maize polyubiquitin promoter and nopaline synthase terminator and introducing the cassette into sugarcane callus by biolistic transformaton.

Levansucrase will probably require apoplastic or vacuolar targeting to maximize access to substrate for conversion.

Fructan may then be recovered from sugarcane juice by ethanol precipitation followed by vacuum-drying.

EXAMPLE 17

Lactic acid production in Sugarcane

The production of Lactic acid (2-Hydroxypropanoic acid) in sugarcane proceeds with the following steps:

(i) Obtain or clone lactate dehydrogenase (LDH) from a number of sources, such a *Lactobacillus* spp. bacterium.

(ii) Expression of the gene in sugarcane, with any necessary changes to the sequence such as codon preference. It is preferred that the introduced gene is expressed in the cytosol, therefore no targeting is required.

(iii) Regenerate plants and evaluate for lactate (or derivative) production.

Lactic acid build-up may cause deleterious effect on cells. There are several ways by which cells can deal with this. One is to remove the acid either by diffusion or transport. The other is modification of the offending chemical and export into the vacuole. Glycosylation is a major signal for this process and lactic acid possesses two potential glycosylation sites.

Traditionally, lactic acid purification has been a complex chemical process. However, recent advances have simplified this process and made it significantly cheaper. It is anticipated that lactic acid can be removed from the post-crushing millstream without great difficulty or extensive modification of existing structures. It is anticipated that the extraction process will be product dependent.

EXAMPLE 18

Adipic acid production in Sugarcane

Adipic acid may be produced in sugarcane by one of two approaches.

I. Synthesis from *cis, cis*-muconic acid

Adipic acid has been produced in transgenic *E. coli* using the metabolic pathway illustrated below. Three genes were introduced into *E. coli* to produce *cis, cis*-muconic acid that was subsequently purified from the fermentation broth and converted to adipic acid by catalytic hydrogenation (step g, 10 % Pt/C, H_2, 3400 kPa, 25°C). This final step has a 97 % conversion efficiency.

The synthesis of *cis, cis*-muconic acid in sugarcane involves making use of the shikimate pathway. In order to use the shikimate pathway to produce *cis, cis*-muconic acid the following biosynthetic enzymes, or homologs thereof are introduced into sugarcane:

Klebsiella pneumoniae 3-dehydroshikimate dehydratase (aroZ) – enzyme d

3-dehydroshikimate → protocatechuate

Klebsiella pneumoniae protocatechuate decarboxylase (aroY) – enzyme e

Protocatechuate → catechol

Acinetobacter calcoaceticus catechol 1,2-dioxygenase (catA) – enzyme f

Catechol + O_2 → *cis, cis*-muconic acid

Introduction of these genes into sugarcane involves constructing an expression cassette by fusing the genes described above to the maize polyubiquitin promoter and nopaline synthase terminator and introducing the cassette into sugarcane callus by biolistic transformation. Catechol is probably produced in most plants, and therefore, it may be unnecessary to clone additional copies of 3-dehydroshikimate dehydratase or protocatechuate decarboxylase. The cloned gene(s) are plastid-targeted since the shikimate pathway reactions reside in this compartment.

The shikimate pathway executes a central role in plant secondary metabolism. This is one of the most active pathways in plants in terms of carbon flux owing to the fact that it is the source of lignin precursors. This makes it an attractive candidate for metabolic engineering.

II. Synthesis from petroselinic acid

Bio-based adipic acid can be obtained through ozonolysis (O_3) of petroselinic acid (18:1 $\Delta^{6\ cis}$), as shown in Figure 10. The coproduct lauric acid is also a potential source of feedstock for detergent manufacture.

The seed oil of the coriander spice plant contains 80-90% petroselinic acid. A 36 kDa putative acyl-ACP desaturase (Genbank accession no. M93115) has been identified from coriander seed extracts and the corresponding cDNA was able to confer the ability to produce petroselinic acid in tobacco callus (Cahoon 1992). Petroselinic acid was quantified from extracted calli by gas chromatography and GC-MS (to determine double bond

position). Tobacco does not normally produce petroselinic acid hence the successful expression of the cDNA in tobacco suggested that this desaturase was sufficient for petroselinic acid formation. This also infers that it may be feasible in sugarcane.

The metabolic pathway for producing petroselinic acid is unclear, however, evidence suggests that it is formed by the desaturation of palmitoyl-ACP by the 36 kDa desaturase followed by elongation to form petroselinic acid (Cahoon 1994).

$$16:0\text{-ACP} \rightarrow 16:1\ \Delta^4\text{-ACP} \rightarrow 18:1\ \Delta^6\text{-ACP}$$

Recent studies have identified a 3-ketoacyl-ACP synthase (Genbank accession no. AF263992) associated with the two-carbon elongation of 16:1 Δ^4-ACP.

Cis, cis-muconic acid in sugarcane juice would be converted to adipic acid by catalytic hydrogenation. The adipic acid in the resultant solution can be recovered by solvent extraction. The solution is contacted with chloroform or methylene chloride and the adipic acid recovered in the aqueous fraction. The aqueous fraction would then be evaporated to yield crystalline adipic acid.

EXAMPLE 19

Production of 1,3-propanediol (1,3-PD) in Sugarcane

1,3-PD is a natural product of glycerol fermentation in a few enterobacteria and clostridia. Fermentation-derived 1,3-PD was not commercially viable for many years due to the high cost of the glycerol feedstock.

The metabolic reactions that convert glycerol to 1,3-PD have been established from *Klebsiella pneumoniae*.

Klebsiella pneumoniae glycerol dehydratase (dhaB)

glycerol → 3-hydroxypropionaldehyde + H_2O

Klebsiella pneumoniae 1,3-propanediol oxidoreductase (dhaT)

3-hydroxypropionaldehyde + NADH → 1,3-propanediol + NAD

Sugarcane does not naturally produce glycerol therefore the reactions that convert triose phosphates to glycerol must also be engineered into sugarcane.

Saccharomyces cerevisiae glycerol-3-phosphate dehydrogenase

dihydroxyacetone phosphate + NADH → glycerol-3-phosphate + NAD

Saccharomyces cerevisiae glycerol-3-phosphatase

glycerol-3-phosphate + ADP → glycerol + ATP

Effectively, all four new genes must be cloned into sugarcane to convert it into a 1,3-PD biofactory. These genes will be assembled into an expression cassette containing the maize polyubiquitin promoter and nopaline synthase terminator. The cassette will be introduced into sugarcane callus by biolistic transformation and expression will be targeted to the cytosol. The accumulation of 1,3-PD in plant tissue will be assayed from plant extracts by conventional HPLC.

1,3-PD can be recovered from sugarcane juice by extraction with cyclohexane followed by vaporization of the residual solvent. Alternatively, distillation may be employed. Use of cyclohexane is environmentally unsound and distillation is energy intensive. Consequently, a method has been patented that describes the use of ion exclusion resins to recover 1,3-PD (see international patent application WO 01/73097).

EXAMPLE 20

Production of 2-phenylethanol (2-PE) in Sugarcane

The production of 2-PE in sugarcane would be achieved in a similar way to previous examples. Briefly, cloned genes for the 2-PE biosynthetic pathway, which has previously been determined, would be obtained. Second, these genes would then be expressed in sugarcane, tailoring the expression pattern and codon usage if required. Finally, a stable product, as a glucose conjugate, is expected.

A biological pathway for the biosynthesis of 2-PE is presented in Figure 11.

The product is produced naturally in roses and hence should not be toxic. Glucosylation of the active group is likely to occur in sugarcane to reduce potential toxicity.

2-PE would be recovered by crushing the cane and refining from juice as for sucrose, and standard production processes for the synthetic form are well established.

2-PE is water-soluble and should be stable for the time normally taken to process sugarcane for sucrose. If sugarcane stores 2-PE as a glucose conjugate then alkaline hydrolysis may be required.

EXAMPLE 21

Characterization of CPL-Expressing and HCHL-Expressing Sugarcane Plants

A chloroplast-targeted version of *E. CPL* situated between the maize ubi-1 promoter and *nos* terminator of the expression construct pU3z-mcs-nos, was co-bombarded with a plasmid containing a selectable marker (pUKN) into embryogenic sugarcane callus to yield the UC series of transgenic lines. The UH series of plants was generated in the same manner using an analogous expression construct that contained the ORF of the *P.*

fluorescens HCHL gene. To serve as controls for the experiments described below, four non-transgenic lines (TC1-TC4) were also regenerated from the same callus material omitting the transformation and selection steps. The regenerated plants were grown in a greenhouse for four weeks and were then analyzed for pHBA accumulation in leaf tissue using HPLC. Only plants that had higher levels than the control plants (46% and 48% of the population for the UC lines and UH lines, respectively) were included in the analysis shown in Figure 2.

Not surprisingly, none of the transgenic plants had significantly higher levels of "free" pHBA than the control plants. Similar to the situation reported for tobacco plants expressing CPL (Siebert *et al.*, *Plant Physiol.* 112: 811-819, 1996) or HCHL (Mayer *et al.*, *Plant Cell* 13:1669-1682, 2001), the only two compounds that accumulated were pHBA glucose conjugates, a phenolic glucoside and a glucose ester. Both compounds contained a single glucose molecule that was attached by a 1-O-□-D linkage to the hydroxyl or carboxyl group of pHBA. The predominant product in all of the plants examined was the phenolic glucoside, which accounted for at least 90% of the pHBA (see below). The mean value for the population was 0.41% ± 0.04% of dry weight (DW), which is almost 30-fold higher than the mean value for the non-transgenic control plants 0.014% ± 0.01% DW. More important, the pHBA glucoside content of the best plant was 1.5% DW, which is equivalent to 0.69% DW free pHBA after correcting for the attached glucose molecule. This value is three times higher than the highest value obtained with transgenic tobacco plants expressing a different chloroplast-targeted version of CPL (Siebert *et al.*, *supra*). The HCHL-expressing sugarcane plants accumulated even higher levels of pHBA. The mean value for total pHBA glucose conjugates in the UH lines was 0.70% ± 0.07% DW, and the highest level observed at this stage of development was 2.6% DW, which is very similar to the best value reported for transgenic tobacco plants expressing the same enzyme (Mayer *et al.*, *supra*).

Based on the results obtained with the 4-week-old plants, a subset of the primary transformants was selected for further evaluation, and leaf levels of pHBA were determined after 16 weeks additional growth (Figure 15). Included in this analysis were

the two CPL-expressing plants that previously exhibited the highest levels of product accumulation (UC63 and UC65) and five HCHL-expressing plants. The methanol-extracted samples were subjected to acid hydrolysis, which quantitatively hydrolyzes both pHBA glucose conjugates, and free pHBA was determined by HPLC.

It was anticipated that pHBA production would continue throughout development and that the 20-week-old plants would have higher levels of pHBA glucosides than the 4-week-old plants. However, the increase in pHBA content with age was not very dramatic nor was it universally observed when product accumulation was expressed on a dry weight basis. Part of the explanation for this is the lower water content of the older plant leaf tissue. For example, the average dry weight to wet weight ratio for the 20-week-old plants was 0.23, while the corresponding value for the 4-week-old plants was 0.15. When this phenomenon is taken into account and product accumulation is expressed on a fresh weight basis it becomes far more apparent that pHBA levels did increase as the plants continued to grow, except for the two CPL-expressing plants.

The 20-week-old primary transformants were large enough to screen for stalk levels of pHBA without damaging the plants. At this stage of development, the oldest stem tissue is semi-mature and new tillers emerge. Since the stalk is the only part of the sugarcane plant that is normally harvested in the existing sugar mill infrastructure, pHBA accumulation in this tissue is the most important gauge for technical success. Leaf and stem samples were taken from 20-week-old plants, and total pHBA was determined by HPLC after methanol extraction and acid hydrolysis. The third internode from the bottom of the plant was the source of stem tissue for this analysis, and the leaf samples were obtained from the third fully unfurled leaf from the top of the plant. Generally speaking, leaf levels of pHBA were considerably higher than stalk levels.

However, the difference was much more pronounced for the CPL-expressing plants. For example, the average stalk to leaf ratio of pHBA for the five UH lines that were examined was 0.324 ± 0.031, and the highest stalk level of pHBA was 0.24% DW, which is equivalent to 0.52% pHBA glucose conjugates. In marked contrast, the corresponding

ratios for UC63 and UC65 were 0.135 and 0.133, respectively, and product accumulation in the stalk of the best plant (UC63) was only 0.06% DW. Since there are no reported values in the literature for pHBA levels in stem tissue for transgenic plants expressing CPL or HCHL, it will be very interesting to see if these observations will extend to other plant systems. Nevertheless, taken together the above results suggest that HCHL is a better catalyst for pHBA production in sugarcane than CPL, and subsequent studies focused on the UH series of plants.

EXAMPLE 22

Localization of pHBA in sugarcane tissue

To gain a better understanding of pHBA accumulation in different parts the plant, leaf and stem segments were sampled from the primary shoot of 20-week-old UH1. The first leaf at the top with a fully visible dewlap was designated "leaf 1" and consecutive leaves down the stalk were numbered in ascending order. The stem segments were numbered similarly with "internode 1" corresponding to the internode immediately above the point of attachment of leaf 1. The results from this analysis are summarized in Figure 15. Note that the values shown refer to total pHBA after acid hydrolysis. Except for the youngest leaf examined, product accumulation in leaves was relatively uniform along the length of the plant achieving a maximum value of ~1.0% DW. Product accumulation also varied along the length of the leaf, with the tip of the leaf having about twice as much pHBA as the base of leaf (data not shown). A similar trend was observed in the stalk, but there was a much larger discrepancy between young stem tissue and old stem tissue. In agreement with the results described above, pHBA levels in mature stem tissue were about 3-fold lower than mature leaf tissue. These results add additional support to the notion that pHBA accumulation in HCHL-expressing sugarcane plants increases as a function of time.

Additional insight on pHBA distribution was obtained from dissection experiments similar to the one shown in Figure 15. The plant that was used for this analysis was 20-week-old UH1. Three different compartments of the stalk were examined: rind, pith, and vascular

bundles. The most pHBA was found in the rind (1% DW), while the pith and vascular bundles had 3- to 4-fold lower levels. Indeed, pHBA levels in the rind were very similar to values obtained from the leaf midrib and leaf lamina.

Of all of the HCHL-expressing primary transformants monitored, UH98 (Fig. 2B) consistently had the highest levels of pHBA in both leaf and stem tissue. When this plant was 20 weeks old pHBA accumulation in leaf tissue was 2.8% DW (leaf lamina, 3.35% DW; leaf midrib, 1.61% DW). The corresponding value for mature stem tissue was 0.67% DW (rind, 0.96% DW; pith, 0.65% DW). Despite these very high levels of pHBA glucose conjugates, UH98 was morphologically indistinguishable from the non-transformed control line TC1 (Fig. 5).

EXAMPLE 23

Construction of cTP-CPL

PCR was used to generate the monocot chloroplast-targeting sequence that was fused to the N-terminus of *E. coli* CPL. The target for amplification was the maize rbcS gene (GenBank accession number Y00322), which codes for the Rubisco small subunit precursor. Primer 1 (5'-CTA CTC ATA ACC **ATG** GCG CCC ACC GTG-3') hybridized to nucleotides 489-505 and introduced a *Nco*I site at the start codon of the transit peptide. Primer 2 (5'-CAT CTT ACT **CAT ATG** CCG CAC CTG CAT GCA CCG GAT CCT TCC G-3') hybridized to nucleotides 616-639 and introduced an *Nde*I site five amino acid residues downstream from the chloroplast cleavage site. The PCR product was cut with *Nco*I and *Nde*I and inserted into pET24a-tTP-CPL (manuscript in preparation), after the latter was cleaved with the same enzymes. pET24a-tTP-CPL contains the gene for a chimeric protein that consists of the tomato Rubisco small subunit transit peptide plus the first four amino acid residues of the 'mature' Rubisco small subunit, fused to the N-terminus of *E. coli* CPL. The plasmid DNA was cut with *Nco*I and *Nde*I to remove the tomato chloroplast-targeting sequence, and this was replaced with PCR-generated maize chloroplast-targeting sequence. The ligation mixture was introduced into *E. coli* DH10B,

and growth was selected on LB media containing kanamycin (50 μg mL^{-1}). A representative plasmid

(pET24a-cTP-CPL) was sequenced and no PCR errors were found. The predicted chloroplast cleavage product of the cTP-CPL fusion protein is a CPL variant with five extra N-terminal amino acid residues (i.e. MQVRH-CPL).

EXAMPLE 24

Generation of CPL and HCHL expression constructs used for sugarcane transformation

The antibiotic selection plasmid pUKN contains the ubi-1 promoter, the neomycin phosphotransferase gene and the *nos* terminator. The plasmid pU3z-mcs-nos was used for cTP-CPL and HCHL expression in sugarcane. This plasmid is a modification of pAHC20 and contains the maize ubi-1 promoter and *nos* terminator. Both genes were inserted in the *Spe*I and *Kpn*I sites of the multicloning region that immediately follows the maize ubi-1 intron. The gene coding for cTP-CPL was amplified from pET24a-cTP-CPL using primers 3 and 4. Primer 3 (5'-CTA CTC ATT **TAC TAG T**CA CCA TGG CGC CCA CCG TGA TG-3') (SEQ ID NO: 50) hybridized to the first 18 nucleotides of the ORF of cTP-CPL and introduced a *Spe*I site upstream from the start codon. Primer 3 also contained a consensus monocot ribosomal binding site, CACC, which is situated between the *Spe*I site and the initiator Met codon. Primer 4 (5'-CAT CTT ACT **GGT ACC** TTT AGT ACA ACG GTG ACG CC-3') (SEQ ID NO: 51) hybridized to the other end of the insert and introduced a *Kpn*I site just after the cTP-CPL stop codon. The PCR product was cut with *Spe*I and *Kpn*I, and ligated into similarly digested pU3z-mcs-nos. The ligation reaction mixture was used to transform *E. coli* DH10B and growth was selected on LB media containing ampicillin (100 μg mL^{-1}). A representative plasmid (pU3z-mcs-nos-cTP-CPL) was sequenced to confirm the absence of PCR errors.

Primers 5 and 6 were used to amplify the *Pseudomonas fluorescens* strain AN103 HCHL gene (GenBank accession number Y13067) from plasmid pFI1039. Primer 5 (5'-CTA

CTC ATT TAC TAG TCA CCA TGA GCA CAT ACG AAG GTC G-3') (SEQ ID NO: 52) hybridized to the first 20 nucleotides of the ORF for HCHL and introduced a unique *Spe*I site upstream from the start codon. Primer 5 also contained a consensus monocot ribosomal binding site (CACC) that is situated between the *Spe*I site and the initiator Met codon. Primer 6 (5'-CAT CTT ACT **GGT ACC** TTC AGC GTT TAT ACG CTT GCA-3') (SEQ ID NO: 53) hybridized to the other end of the insert and introduced a *Kpn*I site just after the HCHL stop codon. The PCR product was cut with *Spe*I and *Kpn*I, and ligated into similarly digested pU3z-mcs-nos. The ligation reaction mixture was used to transform *E. coli* DH10B and growth was selected on LB media containing ampicillin (100 µg mL^{-1}). A representative plasmid (pU3z-mcs-nos-HCHL) was sequenced to confirm the absence of PCR errors.

EXAMPLE 25

Plant transformation

Embryogenic callus from sugarcane cultivar Q117 was prepared essentially as described, and grown in the dark at 27 °C on MS media supplemented with 3 mg L^{-1} of 2,4-dichlorophenoxy acetic acid (2,4-D). The calli were co-transformed with the antibiotic selection plasmid pUKN by microprojectile bombardment Following bombardment and a recovery phase of 2 weeks in the dark, transformants were placed on MS-2,4D selection media supplemented with 60 mg L^{-1} geneticin. Individual callus clumps were maintained separately throughout the selection process. After 6 weeks, antibiotic-resistant calli were transferred to MS media supplemented with geneticin and incubated in the light to initiate plantlet regeneration. At least four plantlets per callus clump were transferred to pots in a glasshouse certified for the physical containment of transgenic plants for further analysis.

EXAMPLE 26

Measurement of accumulated soluble phenolics by HPLC

Soluble phenolics were extracted from 100-200 mg of leaf or stem tissue. The tissue samples was resuspended in 1 mL of 50% v/v methanol and homogenized in a bead beater (Bio101/Savant, Fastprep FP120, Holbrook, New York). The sample were then agitated in an orbital shaker (200 rpm) for 1 hour at 37°C, and clarified by centrifugation. A 550-µL aliquot of extract was transferred to a fresh tube and dried under vacuum, and the dry residue was dissolved in. 200 µL.

When the goal was to convert pHBA and vanillic acid glucose conjugates to free pHBA and vanillic acid an acid hydrolysis step was included. A 200-µL aliquot of the extract was transferred to a fresh tube and dried under vacuum. After adding 200 µL of 1 N HCl to the dry residue and vortexing, the sample was incubated for 2 hours at 95 °C. The sample was then neutralized by adding 200 µL of 1.2 N NaOH.

Soluble phenolics were detected by HPLC at 32 °C using the Novapak C18 column described above. Samples were filtered through 0.2 µm syringe filters and 20 µL of filtrate was injected for each analysis. Mobile phases were the same as previously described. Solvent was pumped at 1 mL min^{-1} using the following gradient conditions: 0 min, 0% B; 80 min, 80% B; 81 min, 100% B; 85 min, 100% B; 86 min, 0% B. Total run time was 90 minutes. An optimized gradient was applied to separate *p*-hydroxybenzoic acid and vanillic acid (0 min, 10% B; 20 min, 50% B; 21 min, 100% B; 24 min, 100% B; 25 min, 10% B; total runtime was 35 minutes). Identified peaks were quantified using authentic standards (Sigma-Aldrich Co.).

Those skilled in the art will appreciate that the invention described herein is susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications. The invention also includes all of the steps, features, compositions and compounds referred to, or indicated in this specification, individually or collectively, and any and all combinations of any two or more of said steps or features.

BIBLIOGRAPHY

Altschul *et al., Nucl. Acids Res. 25:* 3389-3402. 1997;

An, *Methods of Enzymology 153:* 292, 1987;

Ausubel *et al.*, "Current Protocols in Molecular Biology" John Wiley & Sons Inc, 1994-1998, Chapter 15.

Barker *et al., Plant Mol. Biol. 2:* 235-350, 1983;

Bevan *et al., Nucl. Acid. Res. 11:* 369, 1983;

Biebl *et al., Appl. Microbiol. Biotechnol. 52:* 289-297, 1999;

Bohmert *et al., Planta 211:* 841-845, 2000;

Bower *et al., Molec. Breeding 2:* 239-249, 1996;

Braunegg *et al., Eur. J. Appl. Microbiol. Biotechnol. 6:* 29-37, 1978;

Cahoon and Ohlrogge, *Plant physiol. 104:* 827-837, 1994;

Cahoon *et al., Proc. Natl. Acad. Sci. USA 89:* 11184-11188, 1992

Chaumont *et al., PMB 24:* 631-664, 1994;

De Koning, *Can. J. Micro. 41(1):* 303-309, 1995;

Dellaporta *et al.*, in *Chromosome Structure and Function* pp. 263-282, 1988;

Ehrlich, *Proc. Natl. Acad. Sci. USA 75:* 1433, 1978;

Etschmann *et al.; Appl. Microbiol. Biotechnol. 59:* 1-8, 2002;

Fang and Hirsch, *Plant Physiol.* 115:53-68, 1998;

Garfinkel *et al., Cell 27:* 143-153, 1983;

Gerngross, *Nature Biotech. 17:* 541-544, 1999;

Greenspan *et al., J. Cell Biol., 100:* 965-73, 1985;

Greve, *J. Mol. Appl. Genet. 1:* 499-511, 1983;

Ikuta *et al., Biotech 8:* 241, 1990;

Jendrossek *et al., App. Micro. Biotech. 46:* 451-463, 1996;

Karr *et al., Applied and Environmental Microbioilogy 46:* 1339-1344, 1983;

Katz *et al, J. Gen. Microbiol. 129:* 2703, 1983;

Kim *et al., Plant Cell Physiol.* 44(4): 412-414;

Knight *et al., Plant Cell* 11(5): 875-886, 1999;

Lee, *Trends Biotech. 14:* 431-438, 1996;

Mayer *et al., Plant Cell* 13:1669-1682, 2001;

Mekhedov *et al., Plant Mol. Biol. 47:* 507-518, 2001;

Mergaert and Swings, *Indust. Micro. Biotech. 17:* 463-469, 1996;

Niedz *et al., Plant Cell Reports 14:* 403, 1995;

Niu *et al., Biotechnol. Prog. 18*: 201-211, 2002;

Ostle *et al., Appl. Environ. Microbiol. 44:* 238-41, 1982;

Ow *et al., Science 234:* 856, 1986;

Poirier *et al., Science 256:* 520-523, 1992;

Potrykus *et al., Mol. Gene. Genet. 199:* 183, 1985;

Prasad *et al., Appl. Environ. Microbiol.* 69(2): 917-925, 2003;

Prasher *et al., Biochem. Biophys. Res. Comm. 126:* 1259, 1985;

Salomon *et al., EMBO J. 3:* 141-146, 1984;

Sambrook *et al.,* Molecular Cloning, A Laboratory Manual, 2nd Edition, Cold Spring Harbor Press, 1989

Siebert *et al., Plant Physiol.* 112: 811-819, 1996;

Slater *et al., J. Bacteriol. 180:* 667-673, 1998;

Steinbüchel and Schlegel, *Mol. Micro. 5:* 535-542, 1991;

Sticher *et al., Annu. Rev. Phytopathol.* 35: 235-270, 1997;

Sutcliffe, *Proc. Natl. Acad. Sci. USA 75:* 3737, 1978;

Taguchi *et al., FEMS Microbiol. Lett. 198:* 65-71, 2001;

Thillet *et al., J. Biol. Chem. 263:* 12500, 1988;

Veerpoorte and Memelink, *Curr. Opin. Biotech. 13:* 181-187;

Zukowsky *et al., Proc. Natl. Acad. Sci. USA 80:* 1101, 1983;